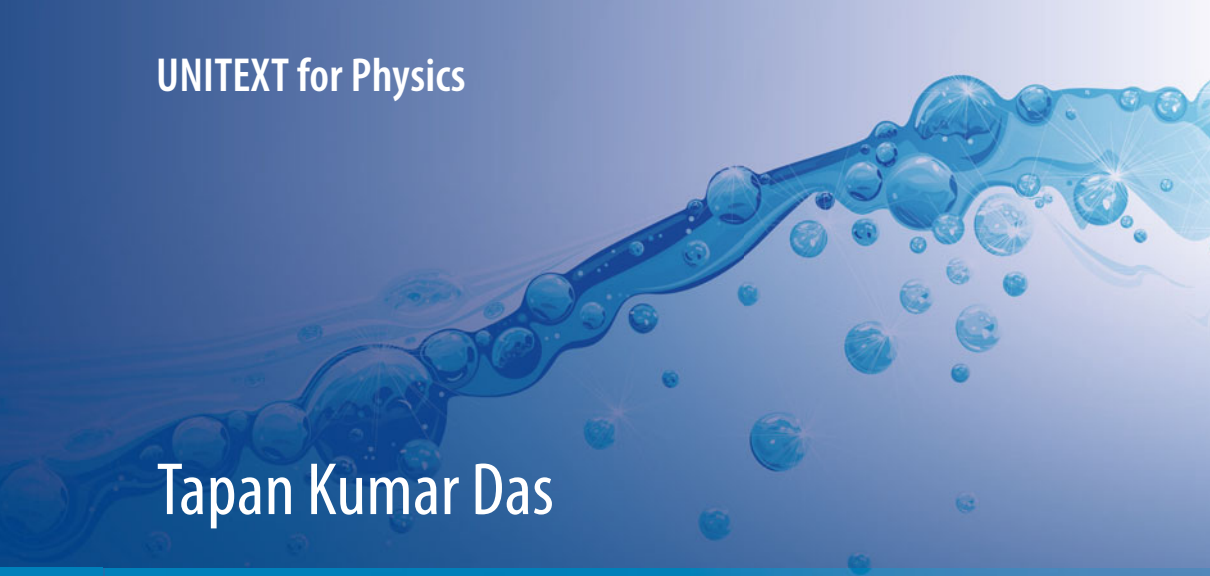
UNITEXT for Physics



Tapan Kumar Das

Quantum Mechanics

Axiomatic Approach and Understanding Through Mathematics

Volume 1

Springer

UNITEXT for Physics

Series Editors

Michele Cini, University of Rome Tor Vergata, Roma, Italy

Stefano Forte, University of Milan, Milan, Italy

Guido Montagna, University of Pavia, Pavia, Italy

Oreste Nicrosini, University of Pavia, Pavia, Italy

Luca Peliti, University of Napoli, Naples, Italy

Alberto Rotondi, Pavia, Italy

Paolo Biscari, Politecnico di Milano, Milan, Italy

Nicola Manini, University of Milan, Milan, Italy

Morten Hjorth-Jensen, University of Oslo, Oslo, Norway

Alessandro De Angelis, Physics and Astronomy, INFN Sezione di Padova, Padova, Italy

UNITEXT for Physics series publishes textbooks in physics and astronomy, characterized by a didactic style and comprehensiveness. The books are addressed to upper-undergraduate and graduate students, but also to scientists and researchers as important resources for their education, knowledge, and teaching.

Tapan Kumar Das

Quantum Mechanics

Axiomatic Approach and Understanding Through Mathematics

Volume 1

Springer

Tapan Kumar Das
Retired Professor
Department of Physics
University of Calcutta
Kolkata, West Bengal, India

ISSN 2198-7882 ISSN 2198-7890 (electronic)
UNITEXT for Physics
ISBN 978-981-99-0493-8 ISBN 978-981-99-0494-5 (eBook)
https://doi.org/10.1007/978-981-99-0494-5

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

To
Rita, my beloved wife

Preface

This book is a comprehensive presentation of immensely popular classroom lectures delivered to graduate students at several universities. There exists a large number of famous textbooks on quantum mechanics. The motivation for adding one more is the following. During my long career of teaching at different universities, I perceived an intrinsic fear in the minds of students encountering quantum mechanics for the first time. The fear stems from an unfounded rumour that the subject is very difficult and not understandable, as it encompasses ideas totally foreign to the common sense. Some textbooks add to this fear by omitting steps in the deduction or failing to explain mathematical ideas behind the theory. In this book, detailed steps in the derivation have been provided to make them easily understandable by the average student. Attempts have been made to introduce mathematical ideas behind new approaches in a simple language. Two chapters (as also a number of sections within a few chapters) detailing the mathematical background have been provided ahead of wherever a new mathematical knowledge is needed for a clear understanding.

Indeed, quantum mechanics encompasses ideas incompatible with our everyday experience arising out of classical physics for macroscopic objects. I found that the way to dispense with the fear of incomprehension is to introduce a new idea starting from our everyday experience. In some cases, it is helpful to give simple analogies with common situations, which make inconceivable ideas conceivable. An example is the classically inconceivable idea that the position and momentum of a particle cannot be simultaneously specified for a microscopic particle. This has been introduced by the following model. Consider a macroscopic particle with a previously known momentum, whose position is determined by "seeing" the particle at a particular position by scattered light (photons). The mass of a photon being negligible compared to the mass of the macroscopic particle, the momentum of the particle remains practically unaffected, so that simultaneous knowledge of the position and momentum of the particle is possible. Now let the particle mass be reduced gradually to that of a microscopic particle. Then the momentum carried away by the photon (which is not measured) becomes comparable with the particle's momentum. This makes the momentum of the microscopic particle appreciably uncertain. Such a transition from the macroscopic to the microscopic world makes sense to the student.

Quantum mechanics thus becomes a theory intimately and inseparably related to *observation, i.e. measurement*. Furthermore, uncertainty principle is first presented in most textbooks as a "principle" introduced by Heisenberg, and much later its mathematical derivation is provided. I found that the mathematical derivation should precede for a better grasp of the situation.

Historically, quantum mechanics gradually took its final shape during the first quarter of the twentieth century, trying to explain the then new experiments involving microscopic systems. These involved several tentative new concepts and *principles* which led initially to the *old quantum theory*. Detailed description of such experiments and the attempted explanations have been left out, since such piecemeal ideas not logically well knitted may lead to misconception and confusion for a beginner. Moreover they become redundant in the current approach. References have been provided to standard textbooks for the interested reader. An axiomatic approach has been adopted in this book, since the whole of quantum mechanics can be developed starting from a few fundamental postulates and physics reasoning only. This has been done in Chaps. 3–5. It is in contrast with the usual approach of presenting the fundamental postulates *after* stating the principles of wave mechanics. An example of solving the Schrödinger equation in the vector space approach has been provided for the standard one-dimensional harmonic oscillator in Chap. 6. It is demonstrated in general (Chap. 8) that bound-state boundary conditions result in discrete energy eigen values, while scattering-state boundary conditions result in continuous energy eigen values. These are then applied to standard topics like solution of Schrödinger equation in coordinate space for a number of simple but important potentials. Chapter 7 provides the basic theory of second order linear differential equations and boundary value problems needed in these treatments. In this connection, the intimate interrelation between physics and mathematics has been emphasized. One notices how mathematics and physics go hand in hand.

The basic idea behind this book is to provide an easy understanding of the principles of quantum mechanics and how they are used in particular situations. Hence I have left out its detailed applications to several more complex problems of physics. Advanced topics have been reserved for the second volume.

I take this opportunity to thank all my students whose intriguing questions helped me shape the lectures, which eventually led to writing this book. I thank my colleagues for helpful discussions during the course of lectures. Special thanks are due to Prof. A. Raychaudhuri and Prof. S. Sengupta, for their valuable help. I also thank my daughter, Arunima, who took a lot of pains to key-in most of the texts together with complicated equations. Last, but not the least, I gratefully remember my wife, Rita, who is no more, for her personal sacrifice in letting me work overtime for this academic pursuit.

Kolkata, India Tapan Kumar Das
August 2022

Contents

1 Introduction 1
1.1 New Experiments and Their Interpretations 4
1.2 Problems 10
References 11

2 Mathematical Preliminary I: Linear Vector Space 13
2.1 Linear Vector Space 13
2.1.1 Formal Definition 14
2.1.2 Subspace 16
2.1.3 Linear Independence of Vectors 17
2.1.4 Basis and Dimension 18
2.2 Scalar (Inner) Product and Inner Product Space 20
2.2.1 Condition of Linear Independence 21
2.2.2 Schwarz Inequality 22
2.2.3 Orthogonality and Normalization 23
2.3 Operators on a Vector Space 25
2.3.1 Eigen Value Equation Satisfied by an Operator 29
2.4 Matrix Representation of Linear Operators 30
2.5 Closure Relation of a Basis 32
2.6 Change of Basis 33
2.7 Dirac's Bra and Ket Notation 36
2.8 Infinite-Dimensional Vector Spaces 39
2.9 Hilbert Space 41
2.10 Problems 44
References 48

3 Axiomatic Approach to Quantum Mechanics 49
3.1 Linear Vector Spaces in Quantum Mechanics 49
3.2 Fundamental Postulates of Quantum Mechanics 50
3.3 Coordinate Space Wave Function: Interpretation 58
3.4 Mathematical Preliminary: Dirac Delta Function 62

3.5 Normalization ... 65
3.6 Problems ... 68
References ... 69

4 Formulation of Quantum Mechanics: Representations and Pictures ... 71
4.1 Position (Coordinate) Representation ... 71
4.2 Momentum Representation ... 74
4.3 Change of Representation ... 77
4.4 Matrix Representation: Matrix Mechanics ... 78
4.5 Math-Prelim: Matrix Eigen Value Equation ... 81
4.6 Quantum Dynamics—Perspectives: Schrödinger, Heisenberg and Interaction Pictures ... 83
4.7 Problems ... 88
References ... 89

5 General Uncertainty Relation ... 91
5.1 Derivation of Uncertainty Relation ... 91
5.2 Minimum Uncertainty Product ... 95
5.3 Problems ... 97
Reference ... 97

6 Harmonic Oscillator: Operator Method ... 99
6.1 Importance of Simple Harmonic Oscillator ... 99
6.2 Energy Eigen Values and Eigen Vectors ... 100
6.3 Matrix Elements ... 105
6.4 Coordinate Space Wave Function ... 106
6.5 Uncertainty Relation ... 108
6.6 Problems ... 110
References ... 110

7 Mathematical Preliminary II: Theory of Second Order Differential Equations ... 111
7.1 Second Order Differential Equations ... 113
7.1.1 Singularities of the Differential Equation ... 114
7.1.2 Linear Dependence of the Solutions ... 115
7.1.3 Series Solution: Frobenius Method ... 117
7.1.4 Boundary Value Problem: Sturm–Liouville Theory ... 122
7.1.5 Connection Between Mathematics and Physics ... 130
7.2 Some Standard Differential Equations ... 134
7.3 Problems ... 146
References ... 147

8 Solution of Schrödinger Equation: Boundary and Continuity Conditions in Coordinate Representation ... 149
8.1 Conditions on Wave Function ... 149
8.2 Eigen Solutions ... 153

8.3 Other Properties . . . 163
8.4 Free Particle Wave Function . . . 169
8.5 Wave Packet and Its Motion . . . 170
8.6 Ehrenfest's Theorem . . . 175
8.7 Problems . . . 178
References . . . 178

9 One-Dimensional Potentials . . . 179
9.1 A Particle in a Rigid Box . . . 179
9.2 A Particle in a Finite Square Well . . . 183
9.3 General Procedure for Bound States . . . 188
9.4 A Particle in a Harmonic Oscillator Well . . . 189
9.5 Wave Packet in a Harmonic Oscillator Well . . . 197
9.6 Potential with a Dirac Delta Function . . . 198
9.7 Quasi-bound State in a δ-Function Barrier . . . 200
9.8 Problems . . . 203
References . . . 204

10 Three-Dimensional Problem: Spherically Symmetric Potential and Orbital Angular Momentum . . . 205
10.1 Connection with Orbital Angular Momentum . . . 205
10.2 Eigen Solution of Orbital Angular Momentum . . . 207
10.3 Radial Equation . . . 213
10.4 Problems . . . 214
References . . . 214

11 Hydrogen-type Atoms: Two Bodies with Mutual Force . . . 215
11.1 Two Mutually Interacting Particles: Reduction to One-Body Schrödinger Equation . . . 215
11.2 Relative Motion of One-Electron H-Type Atoms . . . 217
11.3 Problems . . . 222
References . . . 223

12 Particle in a 3-D Well . . . 225
12.1 Spherically Symmetric Hole with Rigid Walls . . . 225
12.2 Spherically Symmetric Hole with Permeable Walls . . . 228
12.3 A Particle in a Cylindrical Hole with Rigid Walls . . . 230
12.4 3-D Spherically Symmetric Harmonic Oscillator . . . 233
12.5 Problems . . . 237
References . . . 237

13 Scattering in One Dimension . . . 239
13.1 A Free Particle Encountering an Infinitely Rigid Wall . . . 239
13.2 Penetration Through a Finite Square Barrier . . . 241
13.3 Scattering of a Free Particle by a δ-Barrier . . . 245
13.4 Problems . . . 247
Reference . . . 247

14 Scattering in Three Dimension 249
14.1 Kinematics for Scattering 252
14.2 Scattering Cross-Section 254
14.3 Schrödinger Equation 256
14.4 Spherically Symmetric Potential: Method of Partial Waves 260
14.4.1 Optical Theorem 264
14.4.2 Phase Shifts 265
14.4.3 Relation Between Sign of Phase Shift (δ_l) and the Nature (Attractive or Repulsive) of Potential 267
14.4.4 Ramsauer–Townsend Effect 272
14.5 Scattering by a Perfectly Rigid Sphere 274
14.6 Coulomb Scattering 277
14.7 Green's Function in Scattering Theory 281
14.8 Born Approximation 289
14.9 Resonance Scattering 296
14.10 Problems 298
References 299

Appendix: Orthogonality: Physical and Mathematical 301

Index 311

Common Abbreviations Used Throughout This Book

Chap.	Chapter
e.g.	For example
i.e.	That is
LHS	Left hand side
QM	Quantum mechanics
Ref.	Reference
RHS	Right hand side
Sec.	Section
vis-à-vis	In relation to
viz.	Namely
w.r.t.	With respect to

Special Notations Used Throughout This Book

An arrow over a symbol: a common vector
A line under a symbol: a column matrix
Two lines under a symbol: a matrix (square or rectangular)
Re $(\cdots)$: real part of $(\cdots)$
Im $(\cdots)$: imaginary part of $(\cdots)$

Chapter 1
Introduction



Abstract This chapter gives a very brief historical account of new experiments and their interpretations at the beginning of the twentieth century, which led to the proposition of old quantum theory and eventually formulation of quantum mechanics.

Keywords Brief history of quantum mechanics · New experiments · Failure of Classical physics · Old quantum theory · Corpuscular nature of light · Phto-electric effect · Planck-Einstein relations · de Broglie relations · Complementarity principle · Heisenberg uncertainty principle · Bohr's model.

Quantum mechanics is the fundamental theory concerning laws of physics for the motion of microscopic systems, such as molecules, atoms, nuclei and sub-atomic particles. Newtonian laws are not adequate for the description of such systems. The microscopic systems are totally outside direct human perception; hence it is not surprising that quantum mechanics involves physical concepts which are not conceivable in terms of our common daily experience. However, we will see that the new mechanics makes sense when we pay attention to the minuteness of the particles – the smallness of their mass, momentum, energy, etc.

Laws of physics are generalization of observations resulting from careful experiments. Until nearly the end of the nineteenth century, experiments involved macroscopic objects. Naturally, the resulting laws, known collectively as *classical physics*, are consistent with our everyday experience and common perception. Any new experiment involving such objects could be understood in terms of the laws of classical physics. Hence the collection of all the laws of classical physics was thought to be complete and final. However, toward the end of that century and during the first quarter of the twentieth century refined experiments were devised and carried out, which studied the spectral nature of light and probed the microscopic structure of matter. Results of these experiments were not consistent with laws of classical physics. During this period, it became increasingly clear that the laws of classical physics are not adequate – some are even not valid – for the microscopic systems. A number of brilliant physicists, like Planck, Einstein, Bohr, Schrödinger, Heisenberg, de Broglie, Fermi, Bose, Dirac, etc., formulated a set of new laws of physics, collectively known as *quantum mechanics*, which is valid for the microscopic world.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5_1

Before we briefly review the new experiments and conclusions from them, let us summarize implications of the final results.

Experiments showed that ordinary light, whose wave nature has been established by many experiments over several centuries, behaves under certain conditions as a stream of tiny particles, called *photons*, which obey the laws of quantum mechanics. It was also found that light, more precisely electromagnetic radiation, is absorbed and emitted by black bodies and atoms in multiple of a *quantum of energy*, which is the energy of a single photon. The energy of a photon is proportional to the frequency (ν) of the corresponding light and is given as $h\nu$, where h is a universal constant, called Planck's constant. Experimental value of h is very small in macroscopic units, *viz.* $h = 6.626 \times 10^{-27}$ erg.sec. In addition the photo-electric effect demonstrated that the photon of light having frequency ν carries a momentum $p = \frac{h\nu}{c}$, where c is the speed of light. Thus light has a *dual character* – it behaves sometimes as waves and sometimes as particles under different conditions. It is the small but finite value of h that makes the mechanics for microscopic objects different from that for macroscopic objects.

On the other hand, some experiments showed that microscopic particles of matter, under appropriate conditions, show wave nature. For example, beams of electrons and neutrons are diffracted by crystals, like diffraction of light by gratings. The "wave length" (called de Broglie wave length) of a *particle* of momentum p is given by $\lambda = \frac{h}{p}$. Thus there is a *dual nature for all microscopic entities (behaving sometimes as 'particles', and sometimes as 'waves', depending on the situation)*. This is true for all entities – which are traditionally recognized as waves or as particles.

Experiments further demonstrated that the position and momentum of a microscopic particle *cannot simultaneously be known precisely*, in sharp contrast with classical physics.

Since the microscopic systems are well outside our direct perception, it is no wonder that their nature and the laws obeyed by them are also not compatible with our daily experience. However, all the predictions of quantum mechanics have been exhaustively verified by sophisticated experiments performed ever since, which have established the validity of quantum mechanics.

Apart from experimental verification, a sound theory demands that the laws should be logically consistent and valid for all situations. It is not intuitively acceptable that laws of physics would be different for macroscopic and microscopic systems. In the following, we will see that the quantum mechanics is valid universally, and it approaches classical physics in the appropriate limit.

How do we understand the laws of quantum mechanics, which are totally foreign to our common sense? More specifically, how do we reconcile our knowledge of classical physics, which has been well established over many centuries, with this new physics? We can argue that as we gradually reduce the mass of a particle from a macroscopic scale to a microscopic one, the idea that the position and momentum cannot be simultaneously known precisely, makes sense. The main difference with classical physics is the 'process of observation'. In classical (macroscopic) systems, the state of motion is *independent of the 'process of observation' and the observer* – an observation of the system does not interfere with the motion of the system.

But how do we observe (say) the position of a macroscopic particle? The answer is: by the reflected light. Macroscopically, we do not expect the motion of the particle *disturbed* as a result of light being reflected. But with the new experiments, we know that light carries a momentum (although it is extremely small compared to momenta of macroscopic particles) and the process of reflection (i.e. scattering of the photon) will cause an exchange of momentum.

The macroscopic description (that the system remains undisturbed as its position is observed) is indeed a very good approximation, since exchanged momentum ($\sim 10^{-22}$ cgs unit) is completely insignificant compared with the momentum of a macroscopic object (at least ~ 1 cgs unit).

However, for microscopic systems, this 'insignificant amount' may be quite comparable with the momentum of the microscopic system (which is also extremely small in macroscopic units) being observed. Hence the momentum of a microscopic system (with previously precisely known momentum) will now become uncertain by the exchanged amount, as soon as it is observed. On the other hand, initial momentum of a macroscopic particle being about 10^{22} times larger than the exchanged momentum of the photon, the uncertainty in its momentum is indeed quite insignificant, and for all practical purposes we can simultaneously know both the position and the momentum of a macroscopic particle. The extreme smallness of the Planck's constant h makes classical physics appear different from quantum physics. Indeed in the limit of macroscopic physics, the laws of quantum mechanics go over to the laws of classical physics, which is reassuring. We will see in Chap. 8, Sect. 8.6 that Ehrenfest's theorem creates the bridge between classical mechanics and quantum mechanics.

Incidentally, the above discussion shows that quantum mechanics is not independent of the process of observation. Hence an observation of position of a microscopic system (of previously known momentum) will make its momentum uncertain. Such arguments remove the intrinsic mental hindrance of accepting a new mechanics, which at the outset is conceptually difficult to understand.

Thus quantum mechanics was developed over almost a quarter of a century, trying to explain emerging new experimental results. Detailed descriptions of these can be found in standard texts (Refs. Griffiths 1995; Goswami 2009; Greiner 2004; Merzbacher 1998; Reed 2010; Ritchmayer et al. 1955; Sakurai 2000; Scheck 2007; Schiff 1968; Shankar 2008). Historically, the theory was developed piecewise through the introduction of several 'principles', 'postulates', etc. Only after the complete development, the axiomatic nature of its basic principles and the associated mathematical framework were fully understood. Instead of following the historical steps in the development process, quantum mechanics will be presented in this book as an axiomatic theory.

We briefly mention the contents of various chapters. The mathematical background needed for the axiomatic approach appears in Chap. 2. In Chap. 3, we state the fundamental postulates and then all the laws of quantum mechanics follow logically from these and physics logic. The intimate connection of quantum mechanics with measurements and their results are emphasized there. In Chap. 4 we discuss different representations for the same quantum description: position, momentum and

matrix representations and their interrelations. Schrödinger, Heisenberg and interaction pictures are presented for different perspectives in quantum dynamics. Generalized uncertainty relations are derived in Chap. 5 and an example of solution of Schrödinger equationin abstract Hilbert space appears in Chap. 6. Theory of second order differential equation, necessary for solution of the Schrödinger equation, appears in Chap. 7. In Chap. 8, we see in general that bound state boundary conditions result in discrete energy eigen values, while scattering state boundary conditions result in continuous energy eigen values. The procedure is then applied to standard cases in one (Chaps. 9 and 13) and three dimensions (Chaps. 10–12 and 14). Orbital angular momentum appears as an automatic consequence of spherically symmetric three-dimensional cases in Chap. 10. Two-body system with mutual interaction has been shown to reduce to an effective one-body case described by the relative separation vector. This is applied to the hydrogen-type atoms in Chap. 11. Chapter 14 deals with three-dimensional scattering. Analysis of a realistic case requires different and apparently contradictory idealizations between the laboratory setup and the theoretical analysis. These have been made understandable by explaining the widely different length scales in the laboratory and range of interaction, etc.

Although our approach is axiomatic, a very brief description of the important experiments and the associated new ideas is presented in Sec. 1.1 to get a feeling of the historical process. For more details, interested reader is directed to texts of modern physics (e.g. Ref. Ritchmayer et al. 1955).

1.1 New Experiments and Their Interpretations

Corpuscular nature of light

By the end of the nineteenth century, a number of careful measurements involving light showed that their results could not be explained from the knowledge of classical mechanics. Most notably among these is the attempt to understand the spectrum of black body radiation. Attempts to explain the experimental findings required ideas, which are totally incomprehensible according to then existing laws of physics. The most striking among these is Planck's explanation of the spectrum of black body radiation. He postulated in 1900 that thermal energy (an electromagnetic radiation like light) is emitted and absorbed in integral multiples of a single unit, called a 'photon'. This is in stark contrast with the then well established idea that energy is a continuous variable and even an infinitesimal amount of it can be exchanged. Another striking feature is that the energy of a photon is $h\nu$, directly proportional to the frequency (ν) of the radiation, unlike the classical idea that the energy of waves (considering the then prevalent idea of wave nature of light) depends on the intensity of light (i.e. modulus square of the amplitude of oscillation) and not on the frequency. The constant of proportionality (h) in $h\nu$ is called the Planck's constant. Thus energy is 'quantized' and not a continuous variable. Analysis of experimental results showed that the value of the Planck's constant is $h = 6.626 \times 10^{-27}$ erg sec. Note the extremely small value of this constant. The 'quantum of energy' for visi-

ble light, whose frequency is $\nu \sim 10^{15}\,\mathrm{sec}^{-1}$, is as small as 6.6×10^{-12} erg. Thus energy exchange will be in such tiny discrete steps. For an apparatus which is geared for macroscopic measurements, it will appear to be continuous. However when ν is very large, $h\nu$ becomes appreciable and energy exchange cannot be assumed to be continuous. This explains the ultraviolet catastrophe in understanding black body radiation according to classical physics. The corpuscular nature of light got a further support from Einstein's explanation of photo-electric effect in 1905, which is discussed below.

Photo-electric effect
When monochromatic ultraviolet light of frequency ν (which is greater than a cut-off frequency ν_0) falls on a photo-sensitive surface, electrons are emitted with a maximum kinetic energy $\frac{1}{2}mv^2$, where m and v are respectively the mass and maximum speed of the emitted electron. This is known as the photo-electric effect. Classically one expects that this kinetic energy would depend on the intensity of the light. But experimentally it was found to depend not on the intensity, but on the frequency. Einstein adopted Planck's idea and proposed the equation

$$\frac{1}{2}mv^2 = h\nu - W,$$

where $W = h\nu_0$ is the work function, i.e. the binding energy of the electron in the photo-electric material (the energy required to take the electron out of it). Einstein's predictions were verified experimentally.

Compton scattering
The corpuscular nature of light was further supported by Compton scattering of X-rays of frequency ν by atoms. The scattered X-ray had a *lower frequency* ν'. It was also found that $\nu'(\leq \nu)$ depends on the angle of scattering. For classical light waves, the scattered wave should have the same frequency. This phenomenon could be successfully explained assuming X-rays as a stream of photons of energy h times the frequency and momentum $\frac{h}{c}$ times its frequency. Energy and momentum are conserved in the process.

Planck–Einstein relations for photons
Thus by the first decade of the twentieth century, it was established that light has a dual character: it behaves as waves in diffraction and interference experiments, but also behaves as particles (photons) in emission, absorption, scattering and photo-electric effect. Energy and momentum of a photon of frequency ν are given by

$$E = h\nu, \quad p = \frac{h\nu}{c}, \tag{1.1}$$

where c is the speed of light in vacuum. Equations (1.1) are known as Planck–Einstein relations. In this theory the intensity of light is proportional to the flux of photons.

Matter waves: dual character of material particles
In 1924 de Broglie predicted that matter also possesses a dual character: he proposed

that wave length (λ) associated with a particle of momentum $\vec{p}$ is given by *de Broglie relation*:

$$\lambda = \frac{h}{p} = \frac{h}{|\vec{p}|}. \tag{1.2}$$

Note that this is in agreement with the Planck–Einstein relations for the momentum of a photon $p = \frac{h\nu}{c} = \frac{h}{\lambda}$. The prediction of de Broglie was verified by Davisson and Germer and independently by Thompson by observing diffraction of electrons by crystals. Thus electrons behave as waves in the process of diffraction, and as particles when sharp tracks in the Wilson cloud chamber are produced. Hence it became clear that both light and matter possess dual character, sometimes behaving as particles and sometimes behaving as waves, *depending on the experiment, i.e. what is measured.*

To get an idea of the order of magnitude, we see from de Broglie relation, Eq. (1.2), that the wave length of a macroscopic particle of mass 1 g moving at $1\,\mathrm{cm\,s^{-1}}$ is $\sim 10^{-26}$ cm, which is too small to be detected by any macroscopic instrument. On the other hand, when the particle is an electron of mass 9.1×10^{-28} gram moving at 10^7 $\mathrm{cm\,s^{-1}}$, the wave length is $\sim 10^{-6}$ cm, which is comparable to the spacing between scattering centers of the crystal. Hence a diffraction pattern can be observed. The comparison shows that the dual nature of macroscopic particles cannot be detected in macroscopic systems.

Quantized (discrete) values of physical quantities: old quantum theory

Around the same time, several other physical quantities appeared to be 'quantized'. Experimentally observed atomic spectrum contains sharp lines corresponding to specific frequencies. Hence it is discrete i.e. 'quantized'. The discrete spectrum was explained by Bohr's old quantum theory. He postulated that atoms can exist only in 'allowed quantized states' (also called stationary states or stationary orbits), each having a *definite* energy. A radiation 'quantum' of frequency $\nu = (E_2 - E_1)/h$ is emitted when the atom makes a transition from the stationary state having energy E_2 to a lower energy stationary state having energy E_1. The allowed orbits are selected by the condition that the angular momentum must be an integral multiple of $\hbar = h/2\pi$. More generally the *quantization condition for closed orbits* is given by

$$\oint p_i dq_i = n_i h \quad (i = 1, 2, \cdots), \tag{1.3}$$

where p_i is the momentum conjugate to the ith generalized variable q_i and n_i is a positive integer. The integral is to be taken over a complete cycle of the periodic variable (closed orbit). Experimental results for hydrogen-like atoms were in agreement with predictions.

Inadequacy of old quantum theory

However, this simple prescription of the old quantum theory failed for rotational spectra of some diatomic molecules. In addition it cannot be applied to aperiodic systems. This theory also does not provide a proper explanation of the dispersion of light and intensities of spectral lines. Furthermore, it is not clear why the Coulomb force can bind the electron to the nucleus, while the ability to radiate energy by

the charged particle in circular or elliptic orbits is lost in stationary orbits. There is no explanation of the mechanism of emission or absorption of a radiation quantum during the transition from one stationary orbit to another.

In addition, there were intuitive questions about the dual character of light: in a diffraction experiment with a photo-sensitive screen as detector, why the light should behave as waves while traveling, but as particles (photons) when absorbed in the photo-sensitive detector screen? Consider Young's double-slit diffraction experiment. Let us try to explain the *whole experiment* in terms of the corpuscular nature of light. We may consider light as a stream of photons and assume that a photon passing through one slit interferes (interacts) with another passing through the other slit to produce the diffraction pattern. We can reduce the intensity and at the same time increase the exposure time of the photo-sensitive screen. Experiments show that the diffraction pattern remains the same (the diffraction pattern is independent of the intensity of the source). Now in a "thought experiment" we can reduce the intensity to such a value that only one photon is emitted by the source during the time it takes to travel the entire distance. Then how can it interfere with the other photon, which is non-existing? In other words, why should the existence of the second slit, *through which the photon does not pass*, influence its motion?

In order to answer questions of the above type, Bohr introduced the complementarity principle:

Complementarity principle states that microscopic systems cannot be described with the completeness demanded by macroscopic systems.

In other words, for a microscopic system we can talk precisely about a particular property, *only if that property* is measured. Thus when interference of light beams is concerned, light behaves as waves. But when detection by absorption is considered, light behaves as a stream of photons. For the double-slit experiment, we can talk about which slit the photon passes through, *only if we try to determine the transverse position of the photon with respect to the slits*. A careful analysis shows that if this is done, the diffraction pattern (which essentially measures the transverse component of photon momentum) is affected (for details see Ref. Schiff 1968). In fact, if we ascertain ourselves through which slit the photon actually passes, the diffraction pattern would be lost (assuming the validity of the uncertainty principle, see below)! This leads to the idea that precise values of position and momentum of a microscopic system cannot be measured simultaneously. Heisenberg introduced the *uncertainty relations* to account for this.

Heisenberg's uncertainty principle states that the product of uncertainties in the knowledge of a pair of conjugate variables must be at least of the order of $\hbar (= \frac{h}{2\pi})$

$$\begin{aligned} \Delta x \Delta p_x &\gtrsim \hbar \\ \Delta y \Delta p_y &\gtrsim \hbar \\ \Delta z \Delta p_z &\gtrsim \hbar \\ &\text{etc,} \end{aligned} \tag{1.4}$$

where Δx is the uncertainty in the knowledge of the value of x, etc. The uncertainty in position of a microscopic particle can be associated with the de Broglie wave length of the particle. Relation (1.4) shows that if we determine the *exact position* of the particle, i.e. $\Delta x = 0$, then we know *nothing* about the momentum in that direction, $\Delta p_x = \infty$.

We will see later (Chap. 5) that the uncertainty relation between x and p_x is a result of the fact that x and p_x *do not commute*. Then these quantities must be replaced by *operators* instead of numbers, since the latter always commute. It turns out that for quantum systems, the commonly used position variable in coordinate representation (x) will be a multiplying operator, while the conjugate momentum (p_x) will be represented by a differential operator $p_x \to -i\hbar\frac{\partial}{\partial x}$. The representation of conjugate variables by operators is called *quantization* (or more precisely *first quantization* to distinguish it from quantization of fields). We will see later that this leads to "special values" called *eigen values* (which are in general discrete), of the variable. Thus physical quantities for quantum systems are also quantized. Some quantization rules are given below:

Quantization (replacement by differential operators) rules

$$E \to i\hbar\frac{\partial}{\partial t}, \quad \vec{p} \to -i\hbar\vec{\nabla}. \tag{1.5}$$

Quantities like E and $\vec{p}$ are operators in quantum mechanics and act on a function of position variables and time (t), called the wave function (see later).

Some of the other famous experiments whose explanation led to adoption of quantized physical quantities are the following. Experiment by Franck and Hertz showed that energy is lost in discrete amounts when electrons collide with atoms. This demonstrated that not only light but material particles also have quantized energy and lose energy in discrete amounts. The experiment of Stern and Gerlach showed that the magnetic moment of a silver atom along an externally applied magnetic field has discrete values. This is again contrary to the classical idea that the magnetic moment (which is proportional to the angular momentum of the charged particle) can have a range of continuous values, depending on the orientation with respect to the externally applied magnetic field. An analysis leads to the result that the projection in a particular direction (say the z-projection) of the intrinsic (spin) angular momentum of an electron can have only two values, the 'up' and 'down' projections. This was referred to as the space quantization.

From the above discussion it is clear that the unexpected behavior of microscopic particles having extremely small mass, energy, momentum, etc. is due to the extremely small, but *finite* value of the Planck's constant (h). Thus the wave length of a macroscopic particle, according to de Broglie relation, is too small to be measured. Consequently, for all macroscopically practical purposes, the macroscopic particle has zero extent, i.e. no uncertainty in position. In the limit $h \to 0$, the uncertainty relations become irrelevant and *both* Δx and Δp_x can vanish, making both position

and its conjugate momentum simultaneously and precisely measurable. These ideas are consistent with Bohr's correspondence principle:

Bohr's correspondence principle states that all quantum results go over to the classical results in the $h \to 0$ limit.

Rules of quantum mechanics in a nutshell

These principles and associated ideas led to the formulation of the quantum mechanics by the first quarter of the twentieth century. We discuss the basic rules in the following, before an axiomatic development is presented in Chap. 3.

Schrödinger proposed a *wave function*, $\psi(\vec{r}, t)$, which is a mathematical function (in general complex) of the position and time, to *represent* the microscopic particle at $\vec{r}$ at time t. Note that $\psi(\vec{r}, t)$ is a mathematical quantity and represents the particle, but it is not the particle. All quantum mechanically permissible informations about the system can be obtained according to prescriptions involving the wave function. He proposed a partial differential equation, called the Schrödinger equation, together with prescribed boundary conditions to determine $\psi(\vec{r}, t)$, to within a constant phase factor. This function determines the *state* of the system. The equation is based on the classical energy-momentum relation $E = \frac{p^2}{2m} + V(\vec{r}, t)$. Each quantity is treated as an operator using the quantization rules $E \to i\hbar\frac{\partial}{\partial t}$, $\vec{p} \to -i\hbar\vec{\nabla}$ and acting on $\psi(\vec{r}, t)$, resulting in the **Schrödinger equation**:

$$i\hbar\frac{\partial\psi(\vec{r}, t)}{\partial t} = \Big[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r}, t)\Big]\psi(\vec{r}, t). \tag{1.6}$$

This equation is for a particle of mass m moving in a potential $V(\vec{r}, t)$. The differential operator within square brackets on right side of eq. (1.6) is called the Hamiltonian (H) of the system. The wave function $\psi(\vec{r}, t)$ must be *well behaved*, i.e. both ψ and its first derivatives must be continuous everywhere (except at an infinite potential step, where the derivative is discontinuous) and ψ must be finite everywhere. Application of these boundary conditions for a time-independent potential allows a restricted set of energies $E = E_n$, called *energy eigen values*. Wave function $\psi_n(\vec{r}, t)$ corresponding to E_n is called the *energy eigen function*. Energy eigen values are discrete (if the particle is restricted in a region of space) or continuous (if the particle is not so restricted). For energy different from any of the eigen values, the *only solution* of the Schrödinger equationsatisfying the boundary conditions will be the *trivial solution, viz.*, $\psi = 0$. This means that the system cannot exist with energy different from one of the eigen values. The probability density of finding the particle at $\vec{r}$ at time t is given by $|\psi(\vec{r}, t)|^2$, where ψ is *normalized*, i.e. $\int |\psi(\vec{r}, t)|^2 d^3r = 1$ (we will follow the convention that *a definite integral without specified limits*, implies an integral over all space of the variable of integration). The set of all eigen functions forms a complete set. Hence any arbitrary wave function can be expanded in the set of eigen functions.

Since the Schrödinger equationis linear and homogeneous in ψ, the latter is normalizable, and a superposition of solutions is also a solution. A particle partially localized in space is represented by a *wave packet*, which is a superposition of wave

functions satisfying the Schrödinger equation. The spatial extent of the packet is of the order of its position uncertainty.

An *observable*, i.e. a measurable quantity (q) is represented by a Hermitian operator (Q), which is in general a differential operator. An average of its measured value, when the system is in the state ψ is given as $\int \psi^* Q \psi d^3r$.

The results obtained from this procedure – even for complex sub-atomic systems – have been verified experimentally. Hence quantum mechanics appears to be the fundamental theory governing all microscopic systems.

Around the same time, Heisenberg proposed the *matrix mechanics*, which involves column vectors to represent the state of the system and square Hermitian matrices to represent observables. The governing equation is a matrix eigen value equation for the energy operator, called the Hamiltonian matrix. This is completely equivalent to the Schrödinger approach. To distinguish the two approaches, the Schrödinger approach is sometimes called *wave mechanics*.

When quantum mechanics was fully developed, the underlying mathematical structure became clear. It is apparent that the theory is based on a number of assumptions or *axioms*, which result from the generalization of observations – in this case for the microscopic systems. All the rules of quantum mechanics follow from the axiomatic approach (Chap. 3). It also becomes clear from this approach that the theory is in perfect harmony with mathematics, as physics and mathematics go hand in hand. We will discuss the mathematical basis in the following chapter, before presenting the axiomatic approach.

1.2 Problems

1. Calculate the de Broglie wave length of a macroscopic ball of mass 1 g moving at 10 cm/s. Can this length be measured by any existing microscope (even an electron microscope has a resolution $\sim 10^{-8}$ cm)?
2. Next calculate the de Broglie wave length of an electron of mass 9.1×10^{-28} g traveling at 10^8 cm/s. Can this be measured in a suitable experimental setup?
3. Suppose a 10 g bullet traveling at 10^4 cm/s is fired in the x-direction and photographed by a synchronized camera placed at a distance of 100 cm from the gun. The camera has a shutter speed of 10^{-4} sec. Estimate Δx, neglecting air friction. Then calculate p_x and Δp_x. Is Δp_x measurable by any sophisticated instrument?
4. Consider a double-slit diffraction experiment with a monochromatic light of frequency $\nu = 10^{15}$ sec^{-1}. Optical axis is taken as x-axis. The slits are along the y axis, separated by a distance 0.1 cm. Calculate p_x and Δp_y. If a screen is placed perpendicular to the x axis, at a distance of 100 cm from the slits, comment on the possibility of fringe formation in light of the uncertainty principle.

References

Goswami, A.: Quantum Mechanics. Overseas Press, New Delhi (2009)
Greiner, W.: Quantum Mechanics: An Introduction, Springer, New Delhi (2004); Quantum Mechanics: Special Chapters. Springer, Berlin, Heidelberg (1998)
Griffiths, D. J.: Introduction to Quantum Mechanics. Prentice Hall Inc., Upper Saddle River, New Jersey (1995)
Merzbacher, E.: Quantum Mechanics, 3rd edn. John Wiley and Sons Inc, New York (1998)
Reed, B.C.: Quantum Mechanics. Jones and Bartlett Publishers, New Delhi (2010)
Ritchmayer, F.K., Kennard, E.H., Lauritsen, T.: Introduction to Modern Physics. McGraw-Hill Book Co Inc., New York (1955)
Sakurai, J.J. (edited Tuan, S.F.): Modern Quantum Mechanics. Addison-Wesley, Delhi (2000). (Second Indian Reprint)
Scheck, F.: Quantum Mechanics. Springer, Berlin, Heidelberg (2007)
Schiff, L.I.: Quantum Mechanics, 3rd edn. McGraw-Hill Book Company Inc., Singapore (1968)
Shankar, R.: Principles of Quantum Mechanics. Springer, New Delhi (2008). (Second Edition Indian Reprint)

Chapter 2
Mathematical Preliminary I: Linear Vector Space

Abstract This chapter introduces the basic mathematical concepts of linear vector space, which is necessary for an axiomatic approach to quantum mechanics. It is provided for those students who are not familiar with it.

Keywords Definition of linear vector space · Linear independence · Inner product · Schwarz inequality · Operators · Eigen values · Matrix representation · Dirac's bra and ket notation · Hilbert space.

The fundamental equation of motion in quantum mechanics is the time-independent Schrödinger equation. Its acceptable solution (at a given time t) is called the wave function (ψ), which represents a quantum system. In general the wave function is expressed as a function of space coordinates, $\psi(x)$ (in one dimension) or $\psi(\vec{r})$ (in three-dimensional space). A general ψ can be written as a linear superposition of (say) energy eigen functions, $\{\psi_n(\vec{r})\}$

$$\psi(\vec{r}) = \sum_n c_n \psi_n(\vec{r}), \tag{2.1}$$

where c_n are constants. The set of energy eigen functions, $\{\psi_n(\vec{r})\}$ may be orthonormalized. Then we can think of ψ as a vector in an abstract multi-dimensional (in general infinite dimensional) space in which each $\psi_m(\vec{r})$ is a unit vector and c_m is the magnitude of the component of ψ in the "ψ_m-direction", in analogy with ordinary three-dimensional vectors.

Following the mathematical definition of linear vector spaces, we will see that the wave functions in quantum mechanics indeed belong to a linear vector space, called Hilbert space. It gives rise to an elegant and powerful formalism for the development of quantum mechanics.

2.1 Linear Vector Space

The concept of a linear vector space is a generalization of the commonly known three-dimensional vector space, in which each point is represented by a vector $\vec{r}$

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5_2

starting from a chosen origin and ending at that point. If the Cartesian coordinates of this point are (x, y, z), then components of $\vec{r}$ are x, y and z respectively, and $\vec{r}$ is denoted by (x, y, z). Two vectors, $\vec{r}_1$ and $\vec{r}_2$ with components (x_1, y_1, z_1) and (x_2, y_2, z_2), can be added to form a single vector, $\vec{r} = \vec{r}_1 + \vec{r}_2$ with components $(x_1 + x_2, y_1 + y_2, z_1 + z_2)$. Thus vector addition is component-wise addition. For these vectors, it is well-known that "addition of vectors" is both commutative [i.e., vector addition is independent of the order in which they are added: $\vec{r}_1 + \vec{r}_2 = \vec{r}_2 + \vec{r}_1$] and associative [i.e., they can be differently "associated": $\vec{r}_1 + (\vec{r}_2 + \vec{r}_3) = (\vec{r}_1 + \vec{r}_2) + \vec{r}_3$]. The "origin" is considered to be a null vector ($\vec{0}$) with zero components $(0, 0, 0)$. Obviously, $\vec{r} + \vec{0} = \vec{r}$. We can also define a "negative vector", $-\vec{r}$, with components $(-x, -y, -z)$ and clearly $\vec{r} + (-\vec{r}) = \vec{0}$. A vector can also be multiplied by a *number* (called *scalar*), including one (1) and zero (0). Such *scalar multiplication* [for which each component is multiplied by the scalar] satisfies obvious properties. We are also familiar with two-dimensional vectors on the (x, y) plane having properties exactly analogous with those of the three-dimensional vectors. These ideas can be generalized to define vector spaces of more than three dimensions, i.e. a multi-dimensional space. Eventually we can extend the definitions to an infinite-dimensional vector space. In an abstract sense, we can also extend the name "vectors" to a set of objects that are not vectors in the usual sense, but satisfy the above mentioned defining properties. For example, we can consider a set of functions of a variable x, *viz.* $\{\psi_1(x), \psi_2(x), \cdots, \psi_N(x)\}$ to form N vectors in an N-dimensional abstract vector space. We call this a *linear vector space*. The adjective "linear" refers to the linear operations (addition and scalar multiplication) that define the vector space. In the following section, we define an abstract linear vector space (LVS).

2.1.1 Formal Definition

Def. of linear vector space (LVS):
We define a *linear vector space* (denoted by $\mathcal{V}$), to be a collection of elements $\{\psi_a, \psi_b, \psi_c, \cdots\}$, called 'vectors', satisfying the following properties:

1. Vector addition
   For every two vectors ψ_a and ψ_b belonging to (denoted by the symbol $\in$) $\mathcal{V}$, there exists another vector $\in \mathcal{V}$ denoted by $\psi_a + \psi_b$ (the process of association of the vector $\psi_a + \psi_b$ to any two vectors $\psi_a, \psi_b \in \mathcal{V}$ is called *vector addition*), subject to

   (a) vector addition is commutative

   $$\psi_a + \psi_b = \psi_b + \psi_a, \quad \text{for all } \psi_\mathrm{a}, \psi_\mathrm{b} \in \mathcal{V}.$$

   (b) vector addition is associative

   $$\psi_a + (\psi_b + \psi_c) = (\psi_a + \psi_b) + \psi_c, \quad \text{for all } \psi_\mathrm{a}, \psi_\mathrm{b}, \psi_\mathrm{c} \in \mathcal{V}.$$

(c) The vector space $\mathcal{V}$ contains a *null vector*, ψ_0 (also denoted by 0), such that

$$\psi_a + \psi_0 = \psi_a, \quad \text{for all } \psi_a \in \mathcal{V}.$$

(d) For every vector $\psi_a \in \mathcal{V}$, there exists another vector $-\psi_a$ (called *additive inverse*), such that

$$\psi_a + (-\psi_a) = \psi_0, \quad \text{for all } \psi_a \in \mathcal{V}.$$

2. Scalar multiplication
For every $\psi_a \in \mathcal{V}$ and any number μ [belonging to a specified set of numbers (called *field*, see below), e.g. real numbers], there exists another vector $\mu\psi_a$ also $\in \mathcal{V}$ (the process of association of the vector $\mu\psi_a$ to the number μ and any vector $\psi_a \in \mathcal{V}$ is called *scalar multiplication*), subject to

(a) $\lambda(\mu\psi_a) = (\lambda\mu)\psi_a$, for any two numbers λ, μ from the field and all $\psi_a \in \mathcal{V}$.
(b) $1.\psi_a = \psi_a$, for all $\psi_a \in \mathcal{V}$.
(c) $(\lambda + \mu)\psi_a = \lambda\psi_a + \mu\psi_a$, for any two numbers λ, μ from the field and all $\psi_a \in \mathcal{V}$.
(d) $\lambda(\psi_a + \psi_b) = \lambda\psi_a + \lambda\psi_b$, for any number λ from the field and any two vectors $\psi_a, \psi_b \in \mathcal{V}$.

The specified set of numbers $\lambda, \mu, \cdots$ is said to be the *field* over which the vector space is defined. If the numbers are real, we have a real linear vector space. If the numbers are complex, we have a complex linear vector space.

Examples:

1. The three-dimensional ordinary vector space over a field of real numbers is an obvious example. We can easily verify that all the defining properties are satisfied. This example gives us a feeling for the general linear vector space. We can imagine extension to higher dimensions.
2. As an example of an unconventional vector space, consider the set of *all* complex n-tuples, $\psi_a \equiv (c_1^{(a)}, c_2^{(a)}, \cdots, c_n^{(a)})$.

   Def. An ***n*-tuple** is:
   an *ordered set* of n numbers of a specified type: $(c_1, c_2, \cdots, c_n)$.

   Then we can verify that the set of all complex n-tuples forms an LVS over a field of complex numbers, where the process of addition and multiplication are those for complex numbers.

   Verify:
   Let $\psi_a = (c_1^{(a)}, c_2^{(a)}, \cdots, c_n^{(a)})$ and $\psi_b = (c_1^{(b)}, c_2^{(b)}, \cdots, c_n^{(b)})$. Then $\psi_a + \psi_b = (c_1^{(a)} + c_1^{(b)}, c_2^{(a)} + c_2^{(b)}, \cdots, c_n^{(a)} + c_n^{(b)})$ also belongs to the set of complex n-tuples. Clearly such addition is both commutative and associative. The null vector is $\psi_0 = (0, 0, \cdots, 0)$ satisfying property 1(c). For every element $\psi = (c_1, c_2, \cdots, c_n)$ there is an additive inverse element $-\psi = (-c_1, -c_2, \cdots, -c_n)$,

satisfying property 1(d). Also for any complex number λ, we have $\lambda\psi = (\lambda c_1, \lambda c_2, \cdots, \lambda c_n)$ as the scalar multiplication. Clearly, it will satisfy all the defining properties 2(a) to 2(d).
Hence the set of all complex n-tuples over a field of complex numbers forms an LVS.

3. The set of *all* real continuous functions of a variable x defined in an interval $[\alpha, \beta]$ over a field of real numbers forms an LVS.

   Verify:
   Addition: If $\psi_a(x)$ and $\psi_b(x)$ are two such functions, then $\psi_a(x) + \psi_b(x)$ is also a real continuous function in $[\alpha, \beta]$ and belongs to the set. Clearly, such addition is commutative and associative. We can define the null function $\psi_0(x) = 0$, which vanishes everywhere in $[\alpha.\beta]$. For every $\psi(x)$, we can have its negative, $-\psi(x)$, such that at every point x in $[\alpha, \beta]$, its value is the negative of the value of $\psi(x)$. Clearly this one also belongs to the set and properties 1(c) and 1(d) are satisfied.
   Scalar multiplication:
   For every *real* number μ, $\mu\psi(x)$ is a real continuous function defined in $[\alpha, \beta]$ and belongs to the set. Clearly the properties 2(a) to 2(d) are satisfied.
   Thus the set of all real continuous functions in a given interval and over a field of real numbers forms an LVS.

4. The set of all complex numbers over a field of complex numbers forms an LVS. We can easily verify that all the defining properties 1(a) to 1(d) and 2(a) to 2(d) are satisfied, since addition and multiplication of complex numbers give complex numbers (a real number can be considered as a complex number with zero imaginary part). The number zero (0) is the null vector. Here the complex numbers are considered both as vectors and scalars! This is an example of a vector space consisting of a single "ray", since all the vectors are generated from a single number (the number 1) by multiplying it with a complex number.

2.1.2 Subspace

If a subset of the vectors $\psi_a \in \mathcal{V}$ (over a specified field) satisfies all the defining properties 1(a) to 1(d) and 2(a) to 2(d), then this subset of vectors forms an LVS, say, $\mathcal{S}$. The LVS $\mathcal{S}$ is called a *subspace of* $\mathcal{V}$.

Examples:

1. In a three-dimensional ordinary vector space (over a field of real numbers), all the vectors (including the null vector) which lie in the x-y plane form an LVS according to the defining properties. These vectors form a two-dimensional subspace of the three-dimensional ordinary vector space.

2. In Ex. 4 of LVS, if we take the subset of *all real* numbers over a field of real numbers, then we can easily verify that this subset forms an LVS $\mathcal{S}$ over a field of *real numbers*. This LVS $\mathcal{S}$ is a subspace of the LVS $\mathcal{V}$.

2.1.3 Linear Independence of Vectors

Consider an ordinary three-dimensional (3-D) space. If $\vec{r}_1$ and $\vec{r}_2$ are two mutually perpendicular vectors in this space, we cannot express $\vec{r}_1$ as a numerical multiple of $\vec{r}_2$. We say these two vectors are *independent* or more precisely, *linearly independent* (as we try to express one as a linear multiple of the other). Since at any point in this space we can have *at most three* mutually perpendicular vectors, we see that we can have *at most three* linearly independent vectors in this space. We call this space *three dimensional*. In the following we state some general definitions.

Def. Linear combination:
If $\psi_1, \psi_2, \cdots, \psi_n$ are vectors in an LVS and $c_1, c_2, \cdots, c_n$ are numbers from the field, then $\psi = \sum_{i=1}^{n} c_i \psi_i$ is called a linear combination of the vectors $\psi_1, \psi_2, \cdots, \psi_n$.

Def. Linear dependence:
A set of vectors $\{\psi_1, \psi_2, \cdots, \psi_n\}$ in an LVS is said to be linearly dependent if there exists numbers $c_1, c_2, \cdots, c_n$, not all zeros (at least two should be non-zero), such that

$$c_1\psi_1 + c_2\psi_2 + \cdots + c_n\psi_n = 0. \tag{2.2}$$

Def. Linear independence:
The set of vectors $\{\psi_1, \psi_2, \cdots, \psi_n\}$ is said to be linearly independent, if relation (2.2) is satisfied *if and only if (iff)*

$$c_1 = c_2 = \cdots = c_n = 0. \tag{2.3}$$

We will discuss the condition of linear independence in Sect. 2.2.1.

Examples:

1. The set of mutually orthogonal unit vectors $\{\hat{i}, \hat{j}, \hat{k}\}$ along x, y, z directions in Cartesian 3-D space (let us denote it by $\mathcal{R}_3$) is a linearly independent set, since

$$c_1\hat{i} + c_2\hat{j} + c_3\hat{k} = 0$$

is satisfied only if $c_1 = c_2 = c_3 = 0$, as can be seen by taking scalar product with $\hat{i}$, $\hat{j}$ and $\hat{k}$ respectively.

2. Consider the set of n-tuples

$$\begin{aligned}\psi_1 &= (1, 0, \cdots, 0)\\ \psi_2 &= (0, 1, \cdots, 0)\\ &\cdots\\ \psi_n &= (0, 0, \cdots, 1).\end{aligned}$$

This set is linearly independent, since substitution in eq. (2.2) gives

$$(c_1, c_2, \cdots, c_n) = (0, 0, \cdots, 0)$$

which means $c_1 = c_2 = \cdots = c_n = 0$.

2.1.4 Basis and Dimension

Def. Basis:
If a given set of vectors $\{\psi_1, \psi_2, \cdots, \psi_n\}$ in an LVS $\mathcal{V}$ has the properties

1. The set $\{\psi_1, \psi_2, \cdots, \psi_n\}$ is linearly independent.
2. Every vector $\psi \in \mathcal{V}$ can be expressed as a linear combination of $\{\psi_1, \psi_2, \cdots, \psi_n\}$, i.e.

$$\psi = \sum_{i=1}^{n} c_i \psi_i,$$

then the set $\{\psi_1, \psi_2, \cdots, \psi_n\}$ is said to form a *basis* in $\mathcal{V}$. The set $\{\psi_1, \psi_2, \cdots, \psi_n\}$ is also said to *span* the vector space $\mathcal{V}$.

Def. Dimension of a vector space:
A vector space is said to be n-dimensional, if it has a *finite* basis consisting of n elements.

An LVS with *no finite basis* (i.e. $n \to \infty$) is said to be infinite dimensional.

Examples:

1. In ordinary 3-D space $\mathcal{R}_3$, any vector $\vec{A}$ can be expressed as a linear combination of the unit vectors $\hat{i}$, $\hat{j}$ and $\hat{k}$. Thus $\vec{A} = a_x\hat{i} + a_y\hat{j} + a_z\hat{k}$. Hence the set $\{\hat{i}, \hat{j}, \hat{k}\}$ forms a basis and $\mathcal{R}_3$ is three-dimensional.
2. Consider the vector space $\mathcal{C}_n$ formed by the set of all complex n-tuples over a field of complex numbers. We can choose the following set of vectors as a basis in this space

$$\psi_1 = (1, 0, \cdots, 0)$$
$$\psi_2 = (0, 1, \cdots, 0)$$
$$\cdots$$
$$\psi_n = (0, 0, \cdots, 1).$$

We have seen that this set is linearly independent. Then any vector $\psi = (c_1, c_2, \cdots, c_n)$ in $\mathcal{C}_n$ can be written as

$$\psi = c_1\psi_1 + c_2\psi_2 + \cdots + c_n\psi_n.$$

Hence the set of vectors $\{\psi_1, \psi_2, \cdots, \psi_n\}$ forms a basis in $\mathcal{C}_n$ and it has n elements. Thus the space $\mathcal{C}_n$ is n-dimensional.

Note that the basis in a given vector space is *not unique*, but the *dimension* (which is equal to the *number of elements* in the basis) is *unique*. For example, in $\mathcal{R}_3$ we can choose any three mutually orthogonal unit vectors as basis, but the number of such elements must always be three.

Example:
In the above example of $\mathcal{C}_n$, consider the following set of vectors:

$$\phi_1 = (1, 0, \cdots, 0)$$
$$\phi_2 = (1, 1, \cdots, 0)$$
$$\cdots$$
$$\phi_n = (1, 1, \cdots, 1).$$

We can easily verify that these vectors are linearly independent. This set can be chosen as a new basis in $\mathcal{C}_n$, since any arbitrary vector $\psi = (c_1, c_2, \cdots, c_n)$ in $\mathcal{C}_n$ can be written as a linear combination of the set $\{\phi_1, \phi_2, \cdots, \phi_n\}$:

$$\psi = (c_1, c_2, \cdots, c_n) = c_1'\phi_1 + c_2'\phi_2, \cdots + c_n'\phi_n.$$

Equating the middle and the right most sides, we have

$$c_n' = c_n$$
$$c_{n-1}' + c_n' = c_{n-1} \Rightarrow c_{n-1}' = c_{n-1} - c_n$$
$$c_{n-2}' + c_{n-1}' + c_n' = c_{n-2} \Rightarrow c_{n-2}' = c_{n-2} - c_{n-1}$$
$$\cdots$$

Thus all the unknown coefficients $c_1', c_2', \cdots, c_n'$ can be obtained in terms of the known components of the vector ψ. Hence the set $\{\phi_1, \phi_2, \cdots, \phi_n\}$ can as well be chosen as a basis, and the basis is not unique. Note that there must be n elements in this new basis also (otherwise the values of c_i' $(i = 1, n)$ cannot be obtained uniquely). Hence although the basis in a vector space is not unique, its dimension is.

2.2 Scalar (Inner) Product and Inner Product Space

Def. Scalar (or inner) product of two vectors
For every *ordered pair* of vectors ϕ_1 and ϕ_2 in a complex vector space $\mathcal{V}$, we can associate a complex number denoted by (ϕ_1, ϕ_2) and called the scalar (or inner) product of ϕ_1 and ϕ_2, subject to the following defining properties:

1. $(\phi_1, \phi_2) = (\phi_2, \phi_1)^*$.
2. $(\phi_1, \alpha\phi_2 + \beta\phi_3) = \alpha(\phi_1, \phi_2) + \beta(\phi_1, \phi_3)$, ($\alpha, \beta$ are any two complex numbers).
3. $(\phi, \phi) \geq 0$ for every $\phi \in \mathcal{V}$.
4. $(\phi, \phi) = 0$, if and only if $\phi = 0$, the null vector.

We can verify that
$(\alpha\phi_2 + \beta\phi_3, \phi_1) = \alpha^*(\phi_2, \phi_1) + \beta^*(\phi_3, \phi_1)$,
since

$$\begin{aligned} \text{left side} &= (\phi_1, \alpha\phi_2 + \beta\phi_3)^* \quad \text{(by property 1)} \\ &= \alpha^*(\phi_1, \phi_2)^* + \beta^*(\phi_1, \phi_3)^* \quad \text{(by property 2)} \\ &= \alpha^*(\phi_2, \phi_1) + \beta^*(\phi_3, \phi_1) \quad \text{(by property 1)}. \end{aligned}$$

Examples:

1. Consider the ordinary 3-D vector space $\mathcal{R}_3$ and two real vectors:

$$\vec{A} = a_x\hat{i} + a_y\hat{j} + a_z\hat{k},$$
$$\vec{B} = b_x\hat{i} + b_y\hat{j} + b_z\hat{k}.$$

   We can define the inner product as

$$(\vec{A}, \vec{B}) = a_x^* b_x + a_y^* b_y + a_z^* b_z = a_x b_x + a_y b_y + a_z b_z$$

   since all components are real in $\mathcal{R}_3$. This is the usual "dot product" $\vec{A}.\vec{B}$. Clearly

$$(\vec{A}, \vec{A}) = a_x^2 + a_y^2 + a_z^2 \geq 0$$

   and $(\vec{A}, \vec{A}) = 0$, if and only if $a_x = a_y = a_z = 0$, i.e. $\vec{A} = \vec{0}$, the null vector. Since in $\mathcal{R}_3$ we have vectors with real components over a real number field, all the defining properties are automatically satisfied.
2. Consider two vectors

$$\phi_1 = (c_1, c_2, \cdots, c_n)$$
$$\phi_2 = (c_1', c_2', \cdots, c_n')$$

in the LVS $\mathcal{C}_n$ of complex n-tuples over a field of complex numbers. We can define the inner product as

$$(\phi_1, \phi_2) = c_1^* c_1' + c_2^* c_2' + \cdots + c_n^* c_n'.$$

Then

$$(\phi_2, \phi_1) = {c_1'}^* c_1 + {c_2'}^* c_2 + \cdots + {c_n'}^* c_n = (\phi_1, \phi_2)^*.$$

We can easily verify property 2. Also

$$(\phi_1, \phi_1) = |c_1|^2 + |c_2|^2 + \cdots + |c_n|^2 \geq 0.$$

As (ϕ_1, ϕ_1) is a sum of positive numbers, it can vanish if and only if $c_1 = c_2 = \cdots = c_n = 0$, i.e. ϕ_1 is the null vector. These satisfy properties 3 and 4.

Def. Inner product space:
An LVS in which the scalar (or inner) product is defined as above is called an *inner product space*.

From now on, we will consider the LVS denoted by $\mathcal{V}$ to be an inner product space, unless otherwise stated locally.

2.2.1 Condition of Linear Independence

Suppose in an n-dimensional inner product space $\mathcal{V}$ we have m $(m \leq n)$ vectors $\phi_1, \phi_2, \cdots, \phi_m$. We have to check if this set is linearly dependent or independent. Consider the equation

$$c_1\phi_1 + c_2\phi_2 + \cdots + c_m\phi_m = 0, \tag{2.4}$$

where c_i's $(i = 1, \cdots, m)$ are complex numbers. Take inner product of eq. (2.4) successively with each of ϕ_j with $j = 1, \cdots, m$

$$\begin{aligned}
c_1(\phi_1, \phi_1) + c_2(\phi_1, \phi_2) + \cdots + c_m(\phi_1, \phi_m) &= 0 \\
c_1(\phi_2, \phi_1) + c_2(\phi_2, \phi_2) + \cdots + c_m(\phi_2, \phi_m) &= 0 \\
\cdots &= 0 \\
c_1(\phi_m, \phi_1) + c_2(\phi_m, \phi_2) + \cdots + c_m(\phi_m, \phi_m) &= 0.
\end{aligned}$$

This is a set of m linear homogeneous equations for m unknowns $c_1, c_2, \cdots, c_m$. For a non-trivial solution, the determinant of the coefficients must vanish. This determinant W is called Gram-determinant and is given by

$$W = \begin{vmatrix} (\phi_1, \phi_1) & (\phi_1, \phi_2) & \cdots & (\phi_1, \phi_m) \\ (\phi_2, \phi_1) & (\phi_2, \phi_2) & \cdots & (\phi_2, \phi_m) \\ \cdots & \cdots & \cdots & \cdots \\ (\phi_m, \phi_1) & (\phi_m, \phi_2) & \cdots & (\phi_m, \phi_m) \end{vmatrix}. \tag{2.5}$$

Thus if $W = 0$, there exists a non-trivial solution of Eq. (2.4) and the set is linearly dependent. On the other hand if $W \neq 0$, then *only the trivial solution* $c_1 = c_2 = \cdots = c_m = 0$ is possible and the set is linearly independent.

In an n-dimensional vector space, we can also have $2, 3, \cdots, n$ linearly independent vectors, For example in $\mathcal{R}_3$, we can have two linearly independent (say $\hat{i}$, and $\hat{j}$) vectors.

2.2.2 *Schwarz Inequality*

If ϕ_1 and ϕ_2 are any two vectors $\in \mathcal{V}$, then

$$(\phi_1, \phi_1)(\phi_2, \phi_2) \geq |(\phi_1, \phi_2)|^2. \tag{2.6}$$

Proof:
Let λ be any complex number. Then by defining properties of inner product

$$\begin{aligned} ((\phi_1 + \lambda\phi_2), (\phi_1 + \lambda\phi_2)) &\geq 0, \quad \text{(by prop. 3)} \\ (\phi_1 + \lambda\phi_2, \phi_1) + \lambda(\phi_1 + \lambda\phi_2, \phi_2) &\geq 0, \quad \text{(by prop. 2)} \\ (\phi_1, \phi_1) + \lambda^*(\phi_2, \phi_1) + \lambda(\phi_1, \phi_2) + |\lambda|^2(\phi_2, \phi_2) &\geq 0, \quad \text{(by prop. 1)}. \end{aligned} \tag{2.7}$$

This is true for any λ. Let us choose[1]

$$\lambda = -\frac{(\phi_2, \phi_1)}{(\phi_2, \phi_2)}.$$

Then Eq. (2.7) becomes

$$\begin{aligned} (\phi_1, \phi_1) - \frac{(\phi_1, \phi_2)(\phi_2, \phi_1)}{(\phi_2, \phi_2)} - \frac{(\phi_2, \phi_1)(\phi_1, \phi_2)}{(\phi_2, \phi_2)} + \frac{|(\phi_2, \phi_1)|^2}{(\phi_2, \phi_2)} &\geq 0 \\ \text{i.e.,} \qquad (\phi_1, \phi_1) - \frac{|(\phi_1, \phi_2)|^2}{(\phi_2, \phi_2)} &\geq 0 \\ \text{or,} \qquad (\phi_1, \phi_1)(\phi_2, \phi_2) &\geq |(\phi_1, \phi_2)|^2. \end{aligned}$$

[1] This choice can be thought of as minimizing (by differentiating) left side of Eq. (2.7) with respect to λ^*; however second derivative vanishes and the point becomes a point of inflection. For real λ, the left side becomes a minimum (since $(\phi_2, \phi_2) > 0$) for $\lambda = -\frac{\text{Re}(\phi_2, \phi_1)}{(\phi_2, \phi_2)}$.

This proves the Schwarz inequality.

Physical significance: "length" of an abstract vector and "angle" between two abstract vectors.
Consider two ordinary 3-D vectors $\vec{A}$ and $\vec{B}$ in $\mathcal{R}_3$. The inner product

$$(\vec{A}, \vec{B}) = \vec{A}.\vec{B} = a_x b_x + a_y b_y + a_z b_z = |\vec{A}|\,|\vec{B}| \cos\theta, \tag{2.8}$$

where $|\vec{A}|$ and $|\vec{B}|$ are lengths and θ is the angle between $\vec{A}$ and $\vec{B}$. Since $|\cos\theta|^2 \leq 1$, Eq. (2.8) is consistent with Schwarz inequality. This also shows that in an abstract vector space, we can define the "length" of the vector ϕ_1 as $\sqrt{(\phi_1, \phi_1)}$ and the "angle" θ between two vectors ϕ_1 and ϕ_2 as

$$\cos\theta = \frac{(\phi_1, \phi_2)}{\sqrt{(\phi_1, \phi_1)(\phi_2, \phi_2)}}.$$

2.2.3 *Orthogonality and Normalization*

Def. Norm of a vector:
"Norm" of a vector ϕ (denoted by $\| \phi \|$) is defined to be the positive real number $\sqrt{(\phi, \phi)}$ satisfying (norm is the length of the vector)

1. $\| \phi \| \geq 0$.
2. $\| \phi \| = 0$, if and only if ϕ is the null vector.
3. $\| \alpha\phi \| = |\alpha| \, \| \phi \|$.
4. $\| \phi_1 + \phi_2 \| \leq \| \phi_1 \| + \| \phi_2 \|$.

The last property can be easily proved using the Schwarz inequality.
Def. Orthogonality:
Two vectors ϕ_1 and ϕ_2 are said to be orthogonal if $(\phi_1, \phi_2) = 0$.

Def. Normalized vector:
A vector ϕ is said to be normalized if $(\phi, \phi) = 1$. Any vector ϕ' can be normalized by multiplying it by a real number $\frac{1}{\sqrt{(\phi', \phi')}}$ as

$$\phi = \frac{1}{\sqrt{(\phi', \phi')}} \phi'.$$

Def. Orthonormal set:
The set of vectors $\{\phi_1, \phi_2, \cdots, \phi_n\}$ is said to be *orthonormal* if

$$(\phi_i, \phi_j) = \delta_{ij} \quad \text{for all } \phi_i, \phi_j \in \mathcal{V} \quad \text{i.e.} \quad (i, j = 1, 2, \cdots, n).$$

Here δ_{ij} is called Kronecker delta and has the value

$$\begin{aligned}\delta_{ij} &= 1, \quad \text{if } i = j \ \ (i, j \text{ integers})\\ &= 0, \quad \text{for } i \neq j.\end{aligned}$$

If a set of vectors $\{\phi_1, \phi_2, \cdots, \phi_n\} \in \mathcal{V}$ is linearly independent, then we can always construct a set of n orthonormal vectors $\{\psi_1, \psi_2, \cdots, \psi_n\}$. Such a set is particularly convenient. This can be done in several ways. One popular method is the Gram–Schmidt orthonormalization procedure discussed below.

Gram–Schmidt orthonormalization procedure
Suppose we have a set of vectors $\{\phi_1, \cdots, \phi_n\}$ spanning an n-dimensional LVS $\mathcal{V}$, which are linearly independent, but not orthonormal. We will construct a series of vectors $\phi_1', \cdots, \phi_n'$, such that a particular vector ϕ_i' is orthogonal to each one of the preceding vectors (we name the preceding vectors as $\{\psi_j,\ j = 1, \cdots, (i-1)\}$ after normalizing up to the $(i-1)$th member) in the series. Then we normalize ϕ_i' to get ψ_i. Clearly, this results in the orthonormal set $\{\psi_1, \cdots, \psi_n\}$. Thus let

$$\phi_1' = \phi_1, \ \text{ Then } \ \psi_1 = \frac{\phi_1'}{\sqrt{(\phi_1', \phi_1')}} \ \text{ such that } \ (\psi_1, \psi_1) = 1.$$

Next let

$$\phi_2' = \phi_2 - (\psi_1, \phi_2)\psi_1, \ \text{ such that } \ (\psi_1, \phi_2') = 0.$$

Next normalize ϕ_2' to get ψ_2

$$\psi_2 = \frac{\phi_2'}{\sqrt{(\phi_2', \phi_2')}}.$$

At this stage, ψ_1 and ψ_2 are normalized and orthogonal. Next let

$$\phi_3' = \phi_3 - (\psi_1, \phi_3)\psi_1 - (\psi_2, \phi_3)\psi_2, \ \text{ such that } \ (\psi_1, \phi_3') = (\psi_2, \phi_3') = 0,$$

and normalize it to get ψ_3

$$\psi_3 = \frac{\phi_3'}{\sqrt{(\phi_3', \phi_3')}}.$$

At this stage ψ_1, ψ_2 and ψ_3 are orthonormal. This process can be continued to get the orthonormal set $\{\psi_1, \cdots, \psi_n\}$.

Def. Complete set of vectors:
A set of vectors belonging to an LVS is said to be complete, if *any* vector in that LVS can be expressed as a linear combination of the set of vectors. A set of complete

vectors is said to *span* the LVS. By definition, a basis in a finite-dimensional vector space is automatically complete.

In a finite n-dimensional LVS, any set of n linearly independent vectors can be chosen as a basis. It is convenient to choose a set of n orthonormal vectors (they are automatically linearly independent) as the basis. For a finite-dimensional vector space, this set is automatically *complete*, as any vector belonging to the LVS can be expressed as a linear combination of the orthonormal set. Then this set is called *complete orthonormal set* (CONS) of vectors.

The analogy with the 3-D space $\mathcal{R}_3$ is illustrating. The above is equivalent to choosing any three non-coplanar (hence linearly independent) vectors in $\mathcal{R}_3$ to form a coordinate system. If the three non-coplanar vectors are mutually orthogonal (perpendicular) to each other, then we have a Cartesian system.

We can visualize the situation as $n \to \infty$ for an infinite-dimensional LVS; however we have to pay more attention to the question of completeness. The case of an infinite-dimensional vector space will be discussed later in Sect. 2.8.

2.3 Operators on a Vector Space

Def. Operator:

An operator $\hat{A}$ on an LVS $\mathcal{V}$ is a prescription by which every vector $\phi \in \mathcal{V}$ is mapped into another vector ψ also $\in \mathcal{V}$

$$\psi = \hat{A}\phi. \quad \text{for every } \phi \in \mathcal{V};\ \text{also } \psi \in \mathcal{V}.$$

We will denote an operator by a hat ($\hat{}$) above the symbol.

Def. Linear operator:

The operator $\hat{A}$ is *linear* if

$$\hat{A}(c_1\phi_1 + c_2\phi_2) = c_1\hat{A}\phi_1 + c_2\hat{A}\phi_2 \quad \text{for all } \phi_1, \phi_2 \in \mathcal{V}.$$

and c_1, c_2 any two complex numbers.

Examples:

1. *Unit* (or *identity*) operator ($\hat{1}$) is the operator that takes every vector $\in \mathcal{V}$ into itself:
$$\hat{1}\phi = \phi \quad \text{for all } \phi \in \mathcal{V}.$$

2. *Null* operator ($\hat{0}$) is the operator that takes every vector $\in \mathcal{V}$ into the null vector:
$$\hat{0}\phi = 0 \quad \text{for all } \phi \in \mathcal{V}.$$

3. *Rotations in ordinary 3-D space* ($\mathcal{R}_3$)
Let us consider a Cartesian (x, y, z) system of axes and a vector $\vec{r}$ with components (x, y, z). Next consider a rotation (of *all vectors* in $\mathcal{R}_3$) about the z-axis through an angle θ. The vector $\vec{r}$ goes over to $\vec{r}\,'$

$$\vec{r}\,' = \hat{R}_z(\theta)\vec{r},$$

where $\hat{R}_z(\theta)$ is an operator in $\mathcal{R}_3$ that rotates every vector in this space through θ about the z-axis. The vector $\vec{r}\,'$ has components (x', y', z') in the original Cartesian frame. Elementary calculation gives

$$\begin{aligned} x' &= x\cos\theta - y\sin\theta \\ y' &= x\sin\theta + y\cos\theta \\ z' &= z. \end{aligned}$$

We can easily verify that $\hat{R}_z(\theta)$ is linear. Consider rotation of a vector $\vec{S} = \alpha\vec{r}_1 + \beta\vec{r}_2$, such that $\vec{S}\,' = \hat{R}_z(\theta)\vec{S}$. Then

$$\begin{aligned} S'_x &= S_x\cos\theta - S_y\sin\theta \\ &= (\alpha x_1 + \beta x_2)\cos\theta - (\alpha y_1 + \beta y_2)\sin\theta \\ &= \alpha x'_1 + \beta x'_2. \end{aligned}$$

Similarly,

$$\begin{aligned} S'_y &= \alpha y'_1 + \beta y'_2 \\ S'_z &= S_z = \alpha z'_1 + \beta z'_2. \end{aligned}$$

Hence $\hat{R}_z(\theta)\vec{S} = \hat{R}_z(\theta)(\alpha\vec{r}_1 + \beta\vec{r}_2) = \alpha\hat{R}_z(\theta)\vec{r}_1 + \beta\hat{R}_z(\theta)\vec{r}_2$.
Thus $\hat{R}_z(\theta)$ is linear.

4. *Integral transform operator:*
Consider a vector space of continuous functions of x in the interval $[a, b]$. Suppose the operator $\hat{K}$ transforms every continuous function (say $u(x)$) in the space to another continuous function $v(x)$ through

$$v(x) = \int_a^b K(x, y)u(y)dy,$$

where $K(x, y)$ is continuous in both variables x and y in $[a, b]$. The function $K(x, y)$ is called the *kernel function*. The above relation can be represented by the operator equation

$$v = \hat{K}u.$$

We can verify that the operator $\hat{K}$ is linear.

Def. Non-linear operator:
An operator, which is *not linear* is called a nonlinear operator. An important example of such an operator is the *complex conjugation operator*, $\hat{C}$ in the vector space of complex n-tuples

$$\hat{C}(c_1, c_2, \cdots, c_n) = (c_1^*, c_2^*, \cdots, c_n^*)$$

Then

$$\begin{aligned}\hat{C}(\alpha\phi + \alpha'\phi') &= \hat{C}(\alpha c_1 + \alpha' c_1', \alpha c_2 + \alpha' c_2', \cdots, \alpha c_n + \alpha' c_n') \\ &= (\alpha^* c_1^* + \alpha'^* c_1'^*, \alpha^* c_2^* + \alpha'^* c_2'^*, \cdots, \alpha^* c_n^* + \alpha'^* c_n'^*) \\ &= \alpha^*(c_1^*, c_2^*, \cdots, c_n^*) + \alpha'^*(c_1'^*, c_2'^*, \cdots, c_n'^*) \\ &= \alpha^* \hat{C}\phi + \alpha'^* \hat{C}\phi' \end{aligned} \tag{2.9}$$

Thus $\hat{C}$ is not linear. If property (2.9) is satisfied for every vector $\phi, \phi' \in \mathcal{V}$, then $\hat{C}$ is called *anti-linear*.

Defining properties of operators:

1. *Equality of operators:* Two operators $\hat{A}$ and $\hat{B}$ are said to be equal iff

$$\hat{A}\phi = \hat{B}\phi \text{ for all } \phi \in \mathcal{V}.$$

2. *Sum of operators:* The sum of $\hat{A}$ and $\hat{B}$, denoted by $(\hat{A} + \hat{B})$, satisfies

$$(\hat{A} + \hat{B})\phi = \hat{A}\phi + \hat{B}\phi \text{ for all } \phi \in \mathcal{V}.$$

3. *Product of operators:* The product of $\hat{A}$ and $\hat{B}$, denoted by $\hat{A}\hat{B}$, satisfies

$$\hat{A}\hat{B}\phi = \hat{A}(\hat{B}\phi) \text{ for all } \phi \in \mathcal{V}.$$

Note that the operators act from right to left, i.e., $\hat{B}$ acts *first* on ϕ, then $\hat{A}$ acts on the resultant vector. Also note that in general

$$\hat{B}\hat{A} \neq \hat{A}\hat{B}.$$

As an example, consider $\hat{A} = x$ and $\hat{B} = \frac{\mathrm{d}}{\mathrm{d}x}$ on an LVS of continuous differentiable functions. Let $\phi(x)$ be any vector. Then

$$\hat{A}\phi(x) = x\phi(x), \qquad \hat{B}\phi(x) = \frac{\mathrm{d}\phi(x)}{\mathrm{d}x},$$

and

$$\hat{A}\hat{B}\phi(x) = x\frac{\mathrm{d}\phi(x)}{\mathrm{d}x},$$

$$\hat{B}\hat{A}\phi(x) = \frac{\mathrm{d}}{\mathrm{d}x}(x\phi(x)) = x\frac{\mathrm{d}\phi(x)}{\mathrm{d}x} + \phi(x).$$

Thus, $\hat{B}\hat{A}\phi \neq \hat{A}\hat{B}\phi$ and

$$(\hat{B}\hat{A} - \hat{A}\hat{B})\phi(x) = \phi(x).$$

Properties of sum and product operators:

(a) $\hat{A}(\hat{B} + \hat{C}) = \hat{A}\hat{B} + \hat{A}\hat{C}$.
(b) $(\hat{A} + \hat{B})\hat{C} = \hat{A}\hat{C} + \hat{B}\hat{C}$.
(c) $\hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C}$.

Here $\hat{A}$, $\hat{B}$ and $\hat{C}$ are any three operators on $\mathcal{V}$.

Def. Commutator of two operators
The commutator of $\hat{A}$ and $\hat{B}$ is defined as

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$$

In the above example, $[x, \frac{\mathrm{d}}{\mathrm{d}x}] = -\hat{1}$. We will see later from Eqs. (3.6) and (3.7) as also in Chap. 5 that this commutator is closely related to Heisenberg's famous position-momentum uncertainty relation.

Def. Inverse operator:
Consider the operator equation $\psi = \hat{A}\phi$. Suppose there exists an operator $\hat{B}$ on $\mathcal{V}$, such that for *every* ψ we can get back ϕ by the application of $\hat{B}$ on ψ, i.e. $\phi = \hat{B}\psi$. Then $\hat{B}$ is called the *inverse* of $\hat{A}$ and is denoted by $\hat{A}^{-1}$.
Note that *not every operator* can have an inverse. The operator $\hat{A}$ will have an inverse, if for $\phi_1 \neq \phi_2$, $\hat{A}\phi_1 \neq \hat{A}\phi_2$. Example of an operator that does not have an inverse is the *null* operator: $\hat{0}\phi = 0$, since for $\phi_1 \neq \phi_2$, we have $\hat{0}\phi_1 = \hat{0}\phi_2 = 0$.
An operator $\hat{A}$ has an inverse iff $\hat{A}\phi = 0$ implies $\phi = 0$.

Properties:

(a) $\hat{A}\hat{A}^{-1} = \hat{A}^{-1}\hat{A} = \hat{1}$
(b) If $\hat{A}$ is linear $\hat{A}^{-1}$ is also linear.
(c) $(\hat{A}\hat{B})^{-1} = \hat{B}^{-1}\hat{A}^{-1}$

These can be easily proved using the definition of the inverse operator.

Def. Hermitian adjoint:
The Hermitian adjoint $\hat{A}^{\dagger}$ of a linear operator $\hat{A}$ on a vector space $\mathcal{V}$ is defined through

$$(\phi, \hat{A}^{\dagger}\psi) = (\psi, \hat{A}\phi)^* = (\hat{A}\phi, \psi) \quad \text{for all } \phi, \psi \in \mathcal{V}. \tag{2.10}$$

The Hermitian adjoint satisfies

(a) $(\hat{A}^\dagger)^\dagger = \hat{A}$
(b) $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger \hat{A}^\dagger$
(c) If $\hat{A}$ is linear then $\hat{A}^\dagger$ is also linear.

Def. Hermitian operator:

An operator $\hat{A}$ is said to be Hermitian if $\hat{A}^\dagger = \hat{A}$.

2.3.1 Eigen Value Equation Satisfied by an Operator

Def. Eigen value equation:

If an operator $\hat{A}$ satisfies an equation

$$\hat{A}\ \psi_j = \lambda_j\ \psi_j, \tag{2.11}$$

where λ_j is a number and ψ_j is a vector ($j = 1, 2, \cdots, n$), then this equation is called an *eigen value equation*. The number λ_j is called the *eigen value* corresponding to the *eigen vector* ψ_j. The value λ_j is *eigen* (i.e., special), because for any other value $\lambda \neq$ one of λ_j, Eq. (2.11) is satisfied *only for the trivial solution* the null vector, $\psi = 0$.

We will state the following important theorems, without proofs, which can be found in standard texts (Refs. Arfken1966; Chattopadhyay 2006).

Theorem on eigen value and eigen vectors of a Hermitian operator $\hat{A}$ on $\mathcal{V}$:

If the operator is Hermitian i.e., $\hat{A} = \hat{A}^\dagger$, then

(a) All the eigen values are real.
(b) Eigen vectors belonging to *different eigen values* are orthogonal.
(c) *The set of all eigen vectors* forms a complete set in $\mathcal{V}$.

Proofs of these are similar to those for the differential eigen value equations (see Chap. 7, Sec. 7.1.4). Meaning of completeness has been explained in Sect. 2.2.3. If the set of eigen vectors is the set of eigen functions $\{\psi_n(x)\}$ in the interval $[a, b]$, then completeness of the set means that any other function belonging to the same vector space can be expanded as a linear combination of the set $\{\psi_n(x)\}$. Similarly, if the vector space consists of column matrices (see below), then completeness of eigen column matrices ensures that any column matrix belonging to the same vector space can be expressed as a linear combination of the set of all eigen column matrices.

Note that the eigen vectors can be normalized by multiplying by a constant, which does not affect the homogeneous eigen value equation. This is usually done. Then the set of all eigen vectors forms a **complete othonormal set (CONS)**.

Theorem on commuting operators and their simultaneous eigen vectors

Consider two Hermitian operators $\hat{A}$ and $\hat{B}$. Each will satisfy an eigen value equation similar to Eq. (2.11). In general the eigen vectors for the two operators are different.

An interesting situation in quantum mechanics occurs (see Chap. 3) when a vector $\phi_{i,j}$ becomes the eigen vector for both $\hat{A}$ and $\hat{B}$, corresponding to eigen values α_i and β_j respectively

$$\hat{A}\,\phi_{i,j} = \alpha_i \quad \phi_{i,j}$$
$$\hat{B}\,\phi_{i,j} = \beta_j \quad \phi_{i,j}.$$

Such an eigen vector is called a simultaneous eigen vector of $\hat{A}$ and $\hat{B}$ (it will need two indices, corresponding to different eigen values of $\hat{A}$ and $\hat{B}$). This is possible if $\hat{A}$ and $\hat{B}$ commute, i.e. $[\hat{A}, \hat{B}] = 0$. We state an important theorem without proof (for proof see Refs. Arfken 1966; Chattopadhyay 2006).

(a) If two operators commute, they can have simultaneous eigen vectors.
(b) If two operators have simultaneous eigen vectors, they commute.

2.4 Matrix Representation of Linear Operators

Consider a linear operator $\hat{A}$ on an n-dimensional LVS $\mathcal{V}$, which takes ϕ to ψ

$$\psi = \hat{A}\phi. \tag{2.12}$$

Let $\{\phi_1, \phi_2, \cdots, \phi_n\}$ be an orthonormal basis in $\mathcal{V}$. Express ϕ and ψ in this basis:

$$\phi = \sum_{i=1}^{n} c_i \phi_i,$$
$$\psi = \sum_{i=1}^{n} d_i \phi_i. \tag{2.13}$$

Taking inner product of both these with ϕ_i and using orthonormality of the basis, we have

$$c_i = (\phi_i, \phi), \quad d_i = (\phi_i, \psi). \tag{2.14}$$

Thus we get the "component" of a vector "along ϕ_i" as the inner product of ϕ_i with the vector. Substituting Eq. (2.13) in Eq. (2.12)

$$\sum_{i=1}^{n} d_i \phi_i = \hat{A} \sum_{i=1}^{n} c_i \phi_i = \sum_{i=1}^{n} c_i \hat{A} \phi_i \quad \text{(since } \hat{A} \text{ is linear)}.$$

Next, take inner product with ϕ_j, where j runs from $1, \cdots, n$

$$\sum_{i=1}^{n} d_i(\phi_j, \phi_i) = \sum_{i=1}^{n} c_i(\phi_j, \hat{A}\phi_i).$$

Since the basis is orthonormal, $(\phi_j, \phi_i) = \delta_{ji}$ and we have

$$d_j = \sum_{i=1}^{n} A_{ji} c_i \quad (j = 1, \cdots, n), \tag{2.15}$$

where we defined a complex number A_{ji} (called matrix element) as

$$A_{ji} = (\phi_j, \hat{A}\phi_i). \tag{2.16}$$

Let us introduce n-component *column matrices* (a column matrix will be denoted by a line under its symbol)

$$\underline{d} = \begin{pmatrix} d_1 \\ d_2 \\ \cdots \\ \cdots \\ \cdots \\ d_n \end{pmatrix} \quad \text{and} \quad \underline{c} = \begin{pmatrix} c_1 \\ c_2 \\ \cdots \\ \cdots \\ \cdots \\ c_n \end{pmatrix},$$

and an $n \times n$ square matrix (a square or rectangular matrix will be denoted by *two* lines under its symbol)

$$\underline{\underline{A}} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}.$$

Equation (2.15) can be written in matrix form (using matrix multiplication rules)

$$\underline{d} = \underline{\underline{A}}\ \underline{c}\,. \tag{2.17}$$

This is completely equivalent to Eq. (2.12). $\underline{\underline{A}}$ is called the matrix representation of the operator $\hat{A}$ in the basis $\{\phi_1, \phi_2, \cdots, \phi_n\}$. Similarly column matrices (also called column vectors) $\underline{c}$ and $\underline{d}$ are representations of the vectors ϕ and ψ respectively in the same basis. Note that the matrix representation of vectors and operators are basis dependent for the same LVS. We noted earlier that a basis in a given LVS is not unique. So the matrix representation is also not unique for the same vector or operator in the same LVS.

Properties of matrix representation:

Suppose $\underline{\underline{A}}$ and $\underline{\underline{B}}$ are two $n \times n$ matrices respectively representing operators $\hat{A}$ and $\hat{B}$ on an n-dimensional LVS $\mathcal{V}$, with respect to the basis $\{\phi_1, \phi_2, \cdots, \phi_n\}$. Then the following properties can easily be verified:

1. The operator $(\hat{A} + \hat{B})$ is represented by the matrix $(\underline{\underline{A}} + \underline{\underline{B}})$.
2. The operator product $(\hat{A}\hat{B})$ is represented by the product matrix $(\underline{\underline{A}}\ \underline{\underline{B}})$.
3. The operator $(\hat{A})^{-1}$ (provided it exists), is represented by the matrix $(\underline{\underline{A}})^{-1}$. If the inverse of $\hat{A}$ does not exist, the inverse of its representation matrix $\underline{\underline{A}}$ also does not exist.
4. The Hermitian adjoint $\hat{A}^\dagger$ of an operator $\hat{A}$ is represented by the Hermitian adjoint matrix $\underline{\underline{A}}^\dagger$. Consider an orthonormal basis $\{\phi_1, \phi_2, \cdots, \phi_n\}$. Then by Eq. (2.10)

$$(\phi_i, \hat{A}^\dagger \phi_j) = (\phi_j, \hat{A}\phi_i)^*$$
$$\text{i.e.,} \quad (\underline{\underline{A}}^\dagger)_{\text{ij}} = (\underline{\underline{A}})^*_{ji} \quad \Rightarrow \quad \underline{\underline{A}}^\dagger = \underline{\underline{\tilde{A}}}^*,$$

   where $\underline{\underline{\tilde{A}}}$ is the transpose of the matrix $\underline{\underline{A}}$. This is the definition of Hermitian adjoint of a matrix $\underline{\underline{A}}$.
4. The matrix representing a Hermitian operator is a Hermitian matrix.
5. Suppose the operator $\hat{A}$ satisfies an eigen value equation

$$\hat{A}\psi_j = \lambda_j \psi_j, \quad (j = 1, \cdots, n)$$

   where λ_j is the eigen value and ψ_j is the corresponding eigen vector for $(j = 1, 2, \cdots, n)$ of the operator $\hat{A}$. Then the matrix $\underline{\underline{A}}$ representing the operator $\hat{A}$ satisfies a matrix eigen value equation

$$\underline{\underline{A}}\ \underline{d}_j = \lambda_j\ \underline{d}_j,$$

   where the set of eigen values $\{\lambda_j,\ j = 1, \cdots, n\}$ is the *same set of eigen values*, and the column matrix $\underline{d}_j$ represents the eigen vector ψ_j in the chosen basis. Matrix eigen value equation satisfies the same theorems as satisfied by the abstract operator eigen value equation of Sect. 2.3.1, as also those for the differential eigen value equation (see Chap. 7, Sect. 7.1.4).

2.5 Closure Relation of a Basis

Substituting c_i from Eq. (2.14) into Eq. (2.13), we have

$$\phi = \sum_{i=1}^{n} (\phi_i, \phi)\phi_i. \tag{2.18}$$

Let us define an operator $\hat{S}_I = \sum_{i=1}^{n} \phi_i(\phi_i$ through [the symbol $(\phi_i$ implies taking inner product with ϕ_i of the vector on which it acts]

$$\hat{S}_I = \sum_{i=1}^{n} \phi_i(\phi_i \quad \Rightarrow \quad \hat{S}_I\,\phi = \sum_{i=1}^{n} \phi_i(\phi_i, \phi) = \phi \quad \text{by Eq.(2.18).} \tag{2.19}$$

We can easily verify that $\hat{S}_I$ is linear. Since Eq. (2.19) is valid for all $\phi \in \mathcal{V}$, we identify $\hat{S}_I$ with the identity operator

$$\hat{S}_I = \sum_{i=1}^{n} \phi_i(\phi_i = \hat{1}. \tag{2.20}$$

A convenient way of writing it symmetrically is

$$\hat{S}_I = \sum_{i=1}^{n} \phi_i)(\phi_i = \hat{1}, \tag{2.21}$$

where $\phi_i)$ is the vector ϕ_i in the new notation. Reason for such a notation will become clear in Sect. 2.7 below. However, we will continue to use the notation ϕ_i for the vector. Equation (2.21) is called the *closure relation* of the basis set $\{\phi_1, \phi_2, \cdots, \phi_n\}$.

2.6 Change of Basis

Consider a basis $\{\phi_i,\ i = 1, 2, \cdots, N\}$ in an N-dimensional vector space $\mathcal{V}$. We can regard the vectors ϕ_i as coordinate axes in the abstract space. If the basis is orthonormal, then ϕ_i is the unit vector for the ith coordinate axis. We have seen that a vector space can have more than one basis. Suppose that the N-dimensional space $\mathcal{V}$ has two orthonormal bases $\{\phi_i,\ i = 1, 2, \cdots, N\}$ and $\{\phi_i',\ i = 1, 2, \cdots, N\}$. This corresponds to two different sets of orthogonal coordinate axes. Going from one set to the other then corresponds to a rotation of the coordinate axes, with the origin fixed. Then each member of the primed basis can be expanded in the CONS of the unprimed basis:

$$\phi_i' = \sum_{j=1}^{N} \phi_j S_{ji} \quad (i = 1, \cdots, N). \tag{2.22}$$

The matrix $\underline{\underline{S}}$ defines the change of basis from the unprimed to the primed one. Consider an arbitrary vector ψ. It can be expressed as a linear combination of the basis vectors in each of the bases:

$$\psi = \sum_{i=1}^{N} c_i \phi_i, \tag{2.23}$$

as also

$$\psi = \sum_{i=1}^{N} c_i' \phi_i'. \tag{2.24}$$

Substituting Eq. (2.22) in Eq. (2.24), we have

$$\psi = \sum_{i=1}^{N} c_i' \sum_{j=1}^{N} \phi_j S_{ji} = \sum_{j=1}^{N} \Big(\sum_{i=1}^{N} c_i' S_{ji} \Big) \phi_j. \tag{2.25}$$

Comparing with Eq. (2.23), we have

$$c_j = \sum_{i=1}^{N} S_{ji} c_i'. \tag{2.26}$$

This can be written in matrix form as

$$\underline{c} = \underline{\underline{S}} \ \ \underline{c'} \, . \tag{2.27}$$

It is seen that the transformation (2.22) is reversible, hence $\underline{\underline{S}}^{-1}$ exists.

Next consider an operator $\hat{A}$ on $\mathcal{V}$. The action of $\hat{A}$ on ϕ_i is

$$\hat{A}\, \phi_i = \sum_{j=1}^{N} A_{ji}\, \phi_j \, . \tag{2.28}$$

Note that this is consistent with our definition of the matrix element $A_{ji} = (\phi_j, \hat{A}\phi_i)$. Similarly

$$\hat{A}\, \phi_i' = \sum_{l=1}^{N} A_{li}' \quad \phi_l' = \sum_{l=1}^{N} A_{li}' \sum_{k=1}^{N} \phi_k \quad S_{kl} = \sum_{k=1}^{N} \phi_k \sum_{l=1}^{N} S_{kl} \quad A_{li}', \tag{2.29}$$

where Eq. (2.22) has been used. Note that $A_{li}' = (\phi_l', \hat{A}\phi_i')$ is the matrix element of the operator $\hat{A}$ in the primed basis, i.e., $\underline{\underline{A}}'$ is the matrix representation of the operator $\hat{A}$ in the primed basis. Left side of Eq. (2.29) becomes using Eqs. (2.22) and (2.28)

$$\hat{A} \quad \phi_i' = \hat{A} \sum_{j=1}^{N} \phi_j S_{ji} = \sum_{j=1}^{N} S_{ji} \sum_{k=1}^{N} A_{kj} \phi_k = \sum_{k=1}^{N} \phi_k \sum_{j=1}^{N} A_{kj} S_{ji} \ . \tag{2.30}$$

Comparing Eqs. (2.29) and (2.30)

$$\sum_{l=1}^{N} S_{kl} A'_{li} = \sum_{j=1}^{N} A_{kj} S_{ji} \ . \tag{2.31}$$

This in matrix form is

$$\underline{\underline{S}}\,\underline{\underline{A}}' = \underline{\underline{A}}\,\underline{\underline{S}} \ . \tag{2.32}$$

Hence

$$\underline{\underline{A}}' = \underline{\underline{S}}^{-1}\,\underline{\underline{A}} \quad \underline{\underline{S}} \ . \tag{2.33}$$

Thus, in general, $\underline{\underline{A}}'$ is obtained from $\underline{\underline{A}}$ by a similarity transformation. Since both bases are orthonormal we have

$$(\phi_i, \phi_j) = \delta_{ij} \ , \quad (\phi_i', \phi_j') = \delta_{ij} \ . \tag{2.34}$$

Using this in Eq. (2.22), we have

$$S_{ik} = (\phi_i, \phi_k') \ . \tag{2.35}$$

Substituting Eq. (2.22) in the second of Eq. (2.34), we have

$$\sum_{i=1}^{N} S_{ik}^* S_{il} = \delta_{kl} \tag{2.36}$$

i.e.

$$\underline{\underline{S}}^{\dagger} \quad \underline{\underline{S}} = \underline{\underline{1}} \ . \tag{2.37}$$

Thus for orthonormal bases the transformation matrix $\underline{\underline{S}}$ is a unitary matrix. Using the unitarity condition, Eq. (2.33) can be rewritten to express $\underline{\underline{A}}'$ as

$$\underline{\underline{A}}' = \underline{\underline{S}}^{\dagger}\,\underline{\underline{A}}\,\underline{\underline{S}} \ . \tag{2.38}$$

Thus the matrix $\underline{\underline{A}}'$ is obtained by a unitary transformation of the matrix $\underline{\underline{A}}$.

This transformation can also be looked at in the following alternative fashion. Suppose we have a *fixed* basis $\{\phi_i, \ i = 1, 2, \cdots, N\}$. Consider a unitary operator $\hat{U}$ (since it is unitary, it represents a 'rotation' of all vectors in $\mathcal{V}$, see below) which takes every vector ϕ in $\mathcal{V}$ into the 'rotated' vector ϕ'

$$\phi' = \hat{U}\,\phi.$$

Then for an arbitrary operator $\hat{A}$, the operator $\hat{U}$ will 'rotate' the vector $\hat{A}\phi$ into $\hat{U}(\hat{A}\phi)$. Suppose $\hat{A}'$ does the same effect in the 'rotated' system as $\hat{A}$ does for the 'unrotated' system. Thus we define another operator $\hat{A}'$, such that its action on ϕ' (i.e. $\hat{U}\phi$) is the same as the 'rotated' $\hat{A}\phi$ [i.e., $\hat{U}(\hat{A}\phi)$]:

$$\hat{A}'\phi' = \hat{A}'(\hat{U}\phi) = \hat{U}(\hat{A}\phi).$$

As this is true for all ϕ in $\mathcal{V}$, we have from the second equality

$$\hat{A}'\hat{U} = \hat{U}\hat{A} \quad \Rightarrow \quad \hat{A}' = \hat{U}\hat{A}\hat{U}^{\dagger}, \tag{2.39}$$

since $\hat{U}$ is unitary. This will be consistent with Eq. (2.38) if $\underline{\underline{S}}^{\dagger}$ is the matrix corresponding to the unitary operator $\hat{U}$ (in the fixed basis $\{\phi_i,\ i = 1, 2, \cdots, N\}$).

A unitary transformation corresponds to a 'rotation' in a general sense (since the norm, i.e. the 'length' of a vector and the inner product i.e. the 'angle' between any two vectors, remain unchanged in a unitary transformation). For the change of basis, Eq. (2.22), the 'coordinate axes' were rotated, the corresponding unitary matrix being $\underline{\underline{S}}$. When we keep the basis fixed (i.e. coordinate axes are kept fixed) while a unitary transformation ($\hat{U}$) is applied on *every operator and vector*, it corresponds to rotating *the system*, keeping the *axes fixed*. Naturally, then to make the effects of these to be the same, the two 'rotations' must be in opposite sense, which is reflected in the fact that $\underline{\underline{S}}^{\dagger} = \underline{\underline{S}}^{-1}$ represents an 'inverse rotation'. It corresponds to the matrix representation of $\hat{U}$. We will see later that the operator $\hat{U}$ which rotates a spinless system through an angle φ about the direction $\hat{n}$ is given by $\hat{U} = \exp(-\frac{i}{\hbar}\hat{n}.\hat{\vec{L}}\varphi)$, where $\hat{\vec{L}}$ is the operator for orbital angular momentum. Hence $\underline{\underline{S}}$ (corresponding to $\hat{U}^{\dagger}$) represents a rotation about $\hat{n}$ through an angle $-\varphi$ (i.e., in the opposite direction).

2.7 Dirac's Bra and Ket Notation

Dirac introduced a new notation for the prefactor and postfactor of an inner product, which is very convenient and elegant. We have noticed that the scalar product is *linear* with respect to the postfactor, but *anti-linear* with respect to the prefactor:

$$(\psi_1, \lambda\psi_2) = \lambda(\psi_1, \psi_2)$$
$$(\mu\psi_1, \psi_2) = \mu^*(\psi_1, \psi_2).$$

This follows from the property: $(\psi_1, \psi_2) = (\psi_2, \psi_1)^*$. This asymmetry led Dirac to introduce two vector spaces, which are *different but related to each other*. The prefactor and postfactor belong to the *bra vector space* and *ket vector space* respectively

and the two spaces are said to be *dual* to each other. He introduced the notation $\langle\psi_1|$ and $|\psi_2\rangle$ for vectors in the "bra" and "ket" spaces respectively. Taken by itself, each one of the "ket" and "bra" vector spaces forms a linear vector space satisfying the basic properties: for every $|\psi_1\rangle$ and $|\psi_2\rangle$ there exists a $|\psi_1\rangle + |\psi_2\rangle$, subject to definition of a vector space (see Sect. 2.1)

$$\begin{aligned} |\psi_1\rangle + |\psi_2\rangle &= |\psi_2\rangle + |\psi_1\rangle, \\ |\psi_1\rangle + (|\psi_2\rangle + |\psi_3\rangle) &= (|\psi_1\rangle + |\psi_2\rangle) + |\psi_3\rangle \\ |\psi_1\rangle + |0\rangle &= |\psi_1\rangle, \\ |\lambda\psi_1\rangle &= \lambda|\psi_1\rangle, \quad \text{etc.}, \end{aligned}$$

where $|0\rangle$ is the null vector. Similar relations hold for members of the bra vector space, except that it is anti-linear. These two vector spaces are not independent but are *dual* to each other. For every $|\psi_1\rangle$ in the ket vector space, there corresponds a $\langle\psi_1|$ in the bra vector space and *vice versa*, subject to

$$\begin{aligned} |\psi_1\rangle + |\psi_2\rangle &\leftrightarrow \langle\psi_1| + \langle\psi_2| \\ \lambda|\psi_1\rangle &\leftrightarrow \lambda^*\langle\psi_1|. \end{aligned} \tag{2.40}$$

The connection between the prefactor and postfactor spaces is given by defining the scalar product of a prefactor vector with a postfactor vector through

$$\langle\psi_1|\psi_2\rangle = (\psi_1, \psi_2), \tag{2.41}$$

which gives the connection between the Dirac's notation and the earlier notation. Indeed this form is the origin of the names "bra" and "ket", such that the matrix element of an operator c can be expressed as $\langle$bra$|$c$|$ket$\rangle$, breaking the word "bracket". From Eq. (2.41), and the property of inner product, we have

$$\langle\psi_1|\psi_2\rangle = \langle\psi_2|\psi_1\rangle^*. \tag{2.42}$$

Consider a linear operator $\hat{A}$ on the ket vector space, which takes a ket vector $|\psi_1\rangle$ into another ket vector $|\psi_2\rangle$

$$|\psi_2\rangle = \hat{A}|\psi_1\rangle \tag{2.43}$$

such that

$$\begin{aligned} \hat{A}(|\psi\rangle + |\psi'\rangle) &= \hat{A}|\psi\rangle + \hat{A}|\psi'\rangle \\ \hat{A}(\lambda|\psi\rangle) &= \lambda\hat{A}|\psi\rangle \end{aligned}$$

For every linear operator $\hat{A}$ which takes $|\psi_1\rangle$ into $|\psi_2\rangle$ there exists another linear operator $\hat{\tilde{A}}$, which takes $\langle\psi_1|$ into $\langle\psi_2|$

$$\langle\psi_2| = \langle\psi_1|\hat{\tilde{A}} \tag{2.44}$$

The operators in bra space act from right to left, i.e. *backwards*.

Let us take the inner product of $\langle\psi_3|$ with Eq. (2.43)

$$\langle\psi_3|\psi_2\rangle = \langle\psi_3|\{\hat{A}|\psi_1\rangle\} = \langle\psi_2|\psi_3\rangle^* = [\{\langle\psi_1|\hat{\tilde{A}}\}|\psi_3\rangle]^*, \tag{2.45}$$

in which we have used Eq. (2.44). From the definition (2.10) of Hermitian adjoint of an operator [replacing $\hat{A}$ by $\hat{A}^\dagger$, ϕ by ψ_3 and ψ by ψ_1 in Eq. (2.10)] we have

$$\langle\psi_3|\{\hat{A}|\psi_1\rangle\} = \left[\langle\psi_1|\{\hat{A}^\dagger|\psi_3\rangle\}\right]^*.$$

Compare the right side of this equation with the right side of Eq. (2.45), we have

$$\{\langle\psi_1|\hat{\tilde{A}}\}|\psi_3\rangle = \langle\psi_1|\{\hat{A}^\dagger|\psi_3\rangle\} \tag{2.46}$$

Thus in a scalar product $\hat{\tilde{A}}$ acting to the left is equivalent to $\hat{A}^\dagger$ acting to the right. Hence we can identify $\hat{\tilde{A}}$ with $\hat{A}^\dagger$ and leave out the extra unnecessary brackets ({ and }) i.e., $\hat{\tilde{A}} = \hat{A}^\dagger$. Taking Hermitian adjoint

$$\hat{\tilde{A}}^\dagger = \{\hat{A}^\dagger\}^\dagger = \hat{A}. \tag{2.47}$$

Replacing $\hat{A}$ by $\hat{A}^\dagger$ in Eq. (2.46) and then using Eq. (2.47), we have

$$\langle\psi_1|\{\hat{A}|\psi_3\rangle\} = \{\langle\psi_1|\hat{\tilde{A}}^\dagger\}|\psi_3\rangle = \{\langle\psi_1|\hat{A}\}|\psi_3\rangle \equiv \langle\psi_1|\hat{A}|\psi_3\rangle. \tag{2.48}$$

Thus the operator $\hat{A}$ in the middle can be taken as acting either to the left [third term in Eq. (2.48)] or to the right (first term), whichever is convenient. The quantity is denoted by $\langle\psi_1|\hat{A}|\psi_3\rangle$, dropping the unnecessary { and }. If $\hat{A}$ is Hermitian ($\hat{A}^\dagger = \hat{A}$), we have from Eq. (2.48)

$$\langle\psi_3|\hat{A}|\psi_1\rangle = \langle\psi_1|\hat{A}|\psi_3\rangle^*.$$

The earlier notation and Dirac's bra–ket notation are completely equivalent, although the latter has the advantage of making the relations simple and symmetric.

We can understand the symmetric form (2.21) of the closure relation using the bra–ket notation. Consider an operator $\hat{S}$ in an n dimensional LVS $\mathcal{V}$ spanned by the orthonormal basis $\{|\phi_1\rangle, \cdots, |\phi_n\rangle\}$

$$\hat{S} = \sum_{i=1}^{n} |\phi_i\rangle\langle\phi_i|.$$

Any vector $|\psi\rangle$ in $\mathcal{V}$ can be written as a linear combination of the basis vectors

$$|\psi\rangle = \sum_{i=1}^{n} c_i |\phi_i\rangle,$$

which gives $c_i = \langle\phi_i|\psi\rangle$. Then applying $\hat{S}$ on $|\psi\rangle$

$$\hat{S}|\psi\rangle = \sum_{i=1}^{n} |\phi_i\rangle\langle\phi_i|\psi\rangle = \sum_{i=1}^{n} c_i|\phi_i\rangle = |\psi\rangle.$$

This is true for any $|\psi\rangle \in \mathcal{V}$. Hence $\hat{S} = \hat{1}$ i.e.

$$\sum_{i=1}^{n} |\phi_i\rangle\langle\phi_i| = \hat{1}. \tag{2.49}$$

This is the closure relation in $\mathcal{V}$.

Note that a ket vector followed by a bra vector is an operator, while a bra vector followed by a ket vector is a number (in general complex). The operator $P_i \equiv |\phi_i\rangle\langle\phi_i|$ is called the *projection operator on the ith basis vector*, since $P_i|\psi\rangle = c_i|\phi_i\rangle$.

2.8 Infinite-Dimensional Vector Spaces

So far we discussed finite-dimensional vector spaces. Suppose $\{|\phi_i\rangle, i = 1, 2, \cdots, N\}$ forms an orthonormal basis in an N (finite) dimensional vector space. Then any two vectors $|\psi_1\rangle$ and $|\psi_2\rangle$ can be expanded as

$$|\psi_1\rangle = \sum_{i=1}^{N} c_i^{(1)} |\phi_i\rangle,$$
$$|\psi_2\rangle = \sum_{i=1}^{N} c_i^{(2)} |\phi_i\rangle. \tag{2.50}$$

The inner product of $|\psi_1\rangle$ and $|\psi_2\rangle$ is given by

$$\langle\psi_1|\psi_2\rangle = \sum_{i=1}^{N} [c_i^{(1)}]^* c_i^{(2)}. \tag{2.51}$$

For an infinite-dimensional vector space $N \to \infty$ and we have to worry about the convergence of infinite sums. Furthermore, we have to be sure that the infinite sums of vectors on RHS of Eq. (2.50) do indeed belong to the *same* space (Chattopadhyay 2006). To investigate these, we introduce the following definitions.

Def. Cauchy sequence of numbers:
A sequence of complex numbers $\{z_1, z_2, \cdots, z_n, \cdots\}$ is said to form a Cauchy sequence, if $|z_n - z_m| \to 0$ as $n, m \to \infty$.

Def. Cauchy sequence of vectors:
A sequence of vectors $\{|\phi_1\rangle, |\phi_2\rangle, \cdots, |\phi_n\rangle, \cdots\}$ in a linear vector space is said to form a Cauchy sequence, if the norm $\| |\phi_n\rangle - |\phi_m\rangle \| \to 0$ as $n, m \to \infty$.

This means that the "distance" between nth and mth vectors (or the difference of the nth and mth numbers) becomes progressively smaller as $n, m \to \infty$. Note that the vector $|\phi_n\rangle$ must belong to the space as $n \to \infty$.

Def. Completeness:
A linear vector space is said to be complete, if *every Cauchy sequence of vectors in it converges to a vector within the space*.

We can show that a finite-dimensional vector space is always complete. Consider a Cauchy sequence (CS) of vectors $\{|\psi_1\rangle, |\psi_2\rangle, \cdots, |\psi_n\rangle, \cdots\}$ and an orthonormal basis $\{|\phi_1\rangle, |\phi_2\rangle, \cdots, |\phi_N\rangle\}$, N being the (finite) dimension of the space. Expand each member of the CS

$$|\psi_j\rangle = \sum_{i=1}^{N} c_i^{(j)} |\phi_i\rangle.$$

Then the limiting "distance" between any two members of the CS is

$$\begin{aligned}
\lim_{n,m\to\infty} \| |\psi_n\rangle - |\psi_m\rangle \| &= \lim_{n,m\to\infty} \| \sum_{i=1}^{N} (c_i^{(n)} - c_i^{(m)}) |\phi_i\rangle \| \\
&= \lim_{n,m\to\infty} \Big[\sum_{i=1}^{N} |c_i^{(n)} - c_i^{(m)}|^2 \Big]^{\frac{1}{2}} \\
&= 0.
\end{aligned}$$

The last equality (from the left side) is due to the fact that $\{|\psi_i\rangle\}$ forms a CS. This will be true if $\lim_{n,m\to\infty} |c_i^{(n)} - c_i^{(m)}| \to 0$. Thus the sequence of complex numbers $\{c_i^{(1)}, c_i^{(2)}, \cdots, c_i^{(m)}, \cdots\}$ is a CS. Such a sequence of complex numbers is known to converge. Let it converge to c_i. Now, consider the vector $|\psi\rangle = \sum_{i=1}^{N} c_i |\phi_i\rangle$. Note that it is a vector in the same space. Then

$$\lim_{n\to\infty} \| |\psi\rangle - |\psi_n\rangle \| = \lim_{n\to\infty} \Big[\sum_{i=1}^{N} |c_i - c_i^{(n)}|^2 \Big]^{\frac{1}{2}} = 0,$$

(since $\lim_{n\to\infty} c_i^{(n)} = c_i$). Thus the CS $\{|\psi_n\rangle\}$ converges to a vector $|\psi\rangle$ within the same space. Hence the space is complete.

2.9 Hilbert Space

A Hilbert space is a linear vector space which is complete and an inner product is defined in it.

A Hilbert space may be finite or infinite dimensional. As we have seen, a finite-dimensional inner product space is automatically complete. Hence a finite-dimensional inner product space is a Hilbert space. We will consider two examples of an infinite-dimensional Hilbert space.

Ex. 1 – A vector space consisting of vectors which are an infinite sequence of complex numbers.
Consider the space ($\mathcal{V}$) consisting of elements, each having an infinite sequence of complex numbers:

$$|\phi\rangle = (c_1, c_2, \cdots, c_n, \cdots),$$
$$|\phi'\rangle = (c'_1, c'_2, \cdots, c'_n, \cdots).$$

Clearly, vector addition and scalar multiplication are given by

$$|\phi\rangle + |\phi'\rangle = (c_1 + c'_1, c_2 + c'_2, \cdots, c_n + c'_n, \cdots)$$
$$\text{and} \qquad \lambda|\phi\rangle = (\lambda c_1, \lambda c_2, \cdots, \lambda c_n, \cdots).$$

Hence such elements form a vector space. We can easily verify that other properties are also satisfied. Now we have an identity

$$|c_n + c'_n|^2 + |c_n - c'_n|^2 = 2|c_n|^2 + 2|c'_n|^2.$$

Since $|c_n - c'_n|^2 \geq 0$, we have

$$\sum_{n=1}^{\infty} |c_n + c'_n|^2 \leq 2\sum_{n=1}^{\infty} |c_n|^2 + 2\sum_{n=1}^{\infty} |c'_n|^2.$$

Since $\sum_{n=1}^{\infty} |c_n|^2$ and $\sum_{n=1}^{\infty} |c'_n|^2$ are finite (because both vectors $|\phi\rangle$ and $|\phi'\rangle$ are of finite "lengths", being vectors belonging to a vector space), we see that $\sum_{n=1}^{\infty} |c_n + c'_n|^2$ is also finite. Hence $|\phi\rangle + |\phi'\rangle$ belongs to $\mathcal{V}$. We can similarly verify that $\lambda|\phi\rangle$ also belongs to $\mathcal{V}$. The following set of vectors forms a convenient basis:

$$
\begin{aligned}
|e_1\rangle &= (1, 0, 0, \cdots, 0, \cdots) \\
|e_2\rangle &= (0, 1, 0, \cdots, 0, \cdots) \\
|e_3\rangle &= (0, 0, 1, \cdots, 0, \cdots) \\
&\cdots . \\
|e_n\rangle &= (0, 0, 0, \cdots, 1, \cdots) \\
&\cdots
\end{aligned}
$$

As there is an infinite number of basis vectors, this vector space is infinite dimensional. To prove that $\mathcal{V}$ is a Hilbert space, we have to show that it has an inner product defined in it and it is complete.
We can define an inner product as

$$\langle \phi | \phi' \rangle = \sum_{n=1}^{\infty} c_n^* c_n'. \tag{2.52}$$

In order that the inner product is meaningful, we have to show that the sum on right side converges. Now

$$\sum_{n=1}^{\infty} |c_n^* c_n'| = \sum_{n=1}^{\infty} \frac{1}{2}\{|c_n|^2 + |c_n'|^2 - (|c_n - c_n'|^2)\} \;\leq\; \frac{1}{2}\sum_{n=1}^{\infty} |c_n|^2 + \frac{1}{2}\sum_{n=1}^{\infty} |c_n'|^2.$$

Since each of the terms on the right most side is finite (as $|\phi\rangle$ and $|\phi'\rangle$ belong to $\mathcal{V}$ and therefore are of finite lengths), left most side is also finite. Next we note that

$$\left| \sum_{n=1}^{\infty} c_n^* c_n' \right| \leq \sum_{n=1}^{\infty} |c_n^* c_n'| = \text{finite}.$$

Consequently, the right side of Eq. (2.52) converges. Hence the inner product exists.

To prove *completeness*, consider the CS of vectors $\{\cdots, |\phi^{(n)}\rangle, \cdots, \}$ in $\mathcal{V}$. Hence we have

$$\lim_{n,m\to\infty} \;\| \, |\phi^{(n)}\rangle - |\phi^{(m)}\rangle \, \| \;\to\; 0.$$

This means

$$\lim_{n,m\to\infty} \sum_{i=1}^{\infty} |c_i^{(n)} - c_i^{(m)}|^2 \;\to\; 0.$$

This can be true if, for each i, the sequence of complex numbers $\{\cdots, c_i^{(n)}, \cdots\}$ forms a CS. Such a sequence of complex numbers is known to converge, (say) to c_i :

$$\lim_{n\to\infty} c_i^{(n)} = c_i.$$

Now consider the vector $|\phi\rangle = (c_1, c_2, \cdots, c_n, \cdots)$. Then for a large but finite and fixed N

$$\lim_{n,m\to\infty} \sum_{i=1}^{N} \left|c_i^{(n)} - c_i^{(m)}\right|^2 = \lim_{m\to\infty} \sum_{i=1}^{N} \left|c_i - c_i^{(m)}\right|^2 \;\to\; 0.$$

Since this is true for every N, it follows that

$$\lim_{m\to\infty} \sum_{i=1}^{\infty} \left|c_i - c_i^{(m)}\right|^2 \;\to\; 0$$

$$\text{i.e.} \quad \lim_{m\to\infty} \;\| \,|\phi\rangle - |\phi^{(m)}\rangle\| \;\to\; 0.$$

Thus the CS of vectors $\{\cdots, |\phi^{(m)}\rangle, \cdots\}$ converges to a vector $|\phi\rangle$. For completeness, we have to prove that $|\phi\rangle$ belongs to $\mathcal{V}$. Since $\sum_{i=1}^{\infty} |c_i - c_i^{(m)}|^2$ is finite for any large (but finite) m, the vector $|\phi\rangle - |\phi^{(m)}\rangle$ belongs to $\mathcal{V}$. But $|\phi^{(m)}\rangle$ belongs to $\mathcal{V}$, since it is a member of the CS in $\mathcal{V}$. Hence $(|\phi\rangle - |\phi^{(m)}\rangle) + |\phi^{(m)}\rangle$ i.e. $|\phi\rangle$ also belongs to $\mathcal{V}$. Thus $\mathcal{V}$ is complete and it is a Hilbert space.

Ex. 2 – The vector space $\mathcal{V}$ of all square integrable functions in the interval $[a, b]$. A function $f(x)$ is said to be square integrable in $[a, b]$ if

$$\int_a^b |f(x)|^2 \mathrm{d}x < \infty.$$

Then the sum of two square integrable functions $f(x)$ and $g(x)$ is also square integrable. This follows from

$$\int_a^b |f(x) + g(x)|^2 \mathrm{d}x + \int_a^b |f(x) - g(x)|^2 \mathrm{d}x = 2\left[\int_a^b |f(x)|^2 \mathrm{d}x + \int_a^b |g(x)|^2 \mathrm{d}x\right].$$

Since the second integral on the left side is positive definite, we have

$$\int_a^b |f(x) + g(x)|^2 \mathrm{d}x \;\leq\; 2\left[\int_a^b |f(x)|^2 \mathrm{d}x + \int_a^b |g(x)|^2 \mathrm{d}x\right] \;<\; \infty.$$

It is also clear that if $f(x)$ is square integrable, then $cf(x)$ is also, where c is a finite complex constant. With these properties we can easily see that the set of all square integrable functions form a vector space. We can define the inner product as

$$\langle f|g\rangle = \int_a^b f^*(x) g(x) \mathrm{d}x, \tag{2.53}$$

and the norm as

$$\|f\| = \sqrt{\langle f|f\rangle}. \tag{2.54}$$

However we now need to prove that the integral in Eq. (2.53) is finite. We have

$$|f^*(x)g(x)| = \frac{1}{2}\left[|f(x)|^2 + |g(x)|^2 - |f(x) - g(x)|^2\right] \leq \frac{1}{2}\left[|f(x)|^2 + |g(x)|^2\right].$$

Hence

$$\int_a^b \left|f^*(x)g(x)\right|\mathrm{d}x \leq \frac{1}{2}\left[\int_a^b |f(x)|^2\mathrm{d}x + \int_a^b |g(x)|^2\mathrm{d}x\right] < \infty.$$

Now

$$\left|\int_a^b f^*(x)g(x)\mathrm{d}x\right| \leq \int_a^b \left|f^*(x)g(x)\right|\mathrm{d}x < \infty.$$

Thus the defined scalar product is finite and therefore exists. It can also be shown that the space $\mathcal{V}$ is complete (Chattopadhyay 2006). So, $\mathcal{V}$ is an infinite dimensional Hilbert space.

In particular, we note that the acceptable solutions of the Schrödinger equation thus forms a Hilbert space (see Sect. 4.1, Chap. 4 for coordinate space wave functions.)

2.10 Problems

1. Show that any three non-coplanar ordinary vectors are linearly independent. Also show that there cannot be more than three linearly independent vectors in this space.
2. Check whether each one of the following sets, defined for scalar multiplication over the specified field, forms a linear vector space. If so, find the dimension of the vector space and a possible basis set of vectors:

   (a) Set of all real numbers over a field of real numbers.
   (b) Set of all positive integers over a field of positive integers.
   (c) Set of all real numbers over a field of complex numbers.
   (d) Set of all complex numbers over a field of real numbers.
   (e) Set of all $n \times n$ matrices with complex elements over a field of complex numbers.
   (f) Set of all real solutions of the differential equation $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + n^2y = 0$ over a field of real numbers.
   (g) Set of all real continuous functions defined in the interval $[a, b]$ over a field of real numbers.

3. Show that the alternate set of vectors introduced in the "Example" of Sect. 2.1.4 is linearly independent.
4. Prove that

$$\| \phi_1 + \phi_2 \| \leq \| \phi_1 \| + \| \phi_2 \|.$$

Give a physical interpretation.

5. Calculate the "lengths" of the following vectors, consisting of complex n-tuples over a field of complex numbers. Also find the "angle" between the vectors:

$$\phi_1 = (2 + 3i, 0, 0, \cdots, 0)$$
$$\phi_2 = (1, i, 0, 0, \cdots, 0).$$

6. Check whether the following sets of vectors of 4-tuples of complex numbers over a field of complex numbers form a basis

$$\phi_1 = (1, -1, 0, 0)$$
$$\phi_2 = (1, -1, 0, 1)$$
$$\phi_3 = (i, i, 0, 0)$$
$$\phi_4 = (1, 0, 3, 4)$$

7. Prove that in an n-dimensional vector space, every set of $(n + 1)$ vectors is linearly dependent and no set of $(n - 1)$ vectors can be a basis.
8. Prove that the representation of a vector ψ in a linear vector space in terms of a *given* basis $\{\phi_1, \phi_2, \cdots, \phi_n\}$ is unique.
9. Consider a vector space of continuous functions of x in the interval $[a, b]$. The integral transform operator $\hat{K}$ transforms every continuous function $u(x)$ in the space to another continuous function $v(x)$ through

$$v(x) = \int_a^b K(x, y) u(y) dy,$$

where $K(x, y)$ is continuous in both variables x and y in $[a, b]$. Prove that $\hat{K}$ is linear.

10. Use Eq. (2.10) to prove that

(a)

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger \hat{A}^\dagger.$$

(b)
$$(\hat{A}\hat{B}\cdots\hat{F})^{\dagger} = \hat{F}^{\dagger}\cdots\hat{B}^{\dagger}\hat{A}^{\dagger}.$$

11. Prove that

(a)
$$(\hat{A}\hat{B})^{-1} = \hat{B}^{-1}\hat{A}^{-1}.$$

(b)
$$(\hat{A}\hat{B}\cdots\hat{F})^{-1} = \hat{F}^{-1}\cdots\hat{B}^{-1}\hat{A}^{-1}.$$

12. Show that

(a)
$$[\hat{A}\hat{B}, \hat{C}] = [\hat{A}, \hat{C}]\hat{B} + \hat{A}[\hat{B}, \hat{C}].$$

(b)
$$[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}].$$

13. Show that

(a) $\mathrm{Tr}(\underline{\underline{A}} + \underline{\underline{B}}) = \mathrm{Tr}(\underline{\underline{A}}) + \mathrm{Tr}(\underline{\underline{B}})$
(b) $\mathrm{Tr}(\underline{\underline{A}}.\underline{\underline{B}}) = \mathrm{Tr}(\underline{\underline{B}}.\underline{\underline{A}})$.
(c) $\mathrm{Tr}(\underline{\underline{A}}.\underline{\underline{B}}.\underline{\underline{C}}.\underline{\underline{D}}) = \mathrm{Tr}(\underline{\underline{B}}.\underline{\underline{C}}.\underline{\underline{D}}.\underline{\underline{A}})$.
(d) $\mathrm{Det}(\underline{\underline{A}}.\underline{\underline{B}}) = \mathrm{Det}(\underline{\underline{A}}).\mathrm{Det}(\underline{\underline{B}})$,

where Tr and Det denote trace and determinant of a matrix respectively.

14. Prove that the matrix representation of the operator $\hat{C} = \hat{A}\hat{B}$ in a given basis in an n-dimensional vector space is given by the matrix product $(\underline{\underline{C}}) = (\underline{\underline{A}})(\underline{\underline{B}})$, where $(\underline{\underline{A}})$, $(\underline{\underline{B}})$ and $(\underline{\underline{C}})$ are matrix representations of the operators $\hat{A}$, $\hat{B}$ and $\hat{C}$ respectively in the same basis.
15. Prove that the matrix representation of the operator $\hat{A}^{-1}$ in a given basis is given by the inverse of the matrix $(\underline{\underline{A}})$ representing the operator $\hat{A}$ in the same basis.
16. Prove that the norm or "length" of a vector and the "angle" between two vectors in a linear vector space remain unchanged under a unitary transformation. Hence a unitary transformation corresponds to an overall rotation.
17. Show that
$$\| \phi_1 - \phi_2 \| \geq \quad |(\| \phi_1 \| - \| \phi_2 \|) \quad |,$$
where ϕ_1 and ϕ_2 are any two vectors in a vector space. Give a geometrical interpretation of this result.
18. If $\hat{A}$ and $\hat{B}$ are two operators and $\hat{B}$ is defined through the relation
$$(\psi_a, \hat{A}\psi_b) = (\psi_b, \hat{B}\psi_a),$$
where ψ_a and ψ_b are any two vectors, show that $\hat{B}$ is not linear.

19. In a given representation, if the operator $\hat{A}$ is diagonal with its diagonal elements all different, and $\hat{A}$ and $\hat{B}$ commute, show that $\hat{B}$ is also diagonal in this representation.
20. The functions $f(x)$ and $g(x)$ are defined in the interval $[-1, 1]$ and $f(x)$ is an even function, while $g(x)$ is an odd function. Show that $f(x)$ and $g(x)$ are linearly independent.
21. If P_a is a projection operator defined by $P_a\psi = \psi_a(\psi_a, \psi)$, show that P_a is linear and $P_a^2 = P_a$.
22. The orthonormal set of vectors $\{\phi_1, \phi_2, \phi_3\}$ form a basis in a linear vector space. An operator $\hat{A}$ is such that it transforms ϕ_1 to ϕ_2, ϕ_2 to ϕ_3 and ϕ_3 to ϕ_1.
    (a) Find the matrix representation of the operator $\hat{A}$, in the given basis.
    (b) Is this operator Hermitian? Unitary? What are the eigen values? Obtain corresponding eigen vectors.
23. Suppose $\underline{\underline{A}}$ is an $n \times n$ real symmetric matrix and

$$\underline{\underline{B}} = \mathrm{e}^{\underline{\underline{A}}}.$$

    Use the series expansion

$$\underline{\underline{B}} = \mathrm{e}^{\underline{\underline{A}}} = \underline{\underline{1}} + \underline{\underline{A}} + \frac{\underline{\underline{A}}^2}{2!} + \frac{\underline{\underline{A}}^3}{3!} + \cdots$$

    and apply the unitary transformation $\underline{\underline{U}}$ which diagonalizes $\underline{\underline{A}}$ to prove that

$$\det(\underline{\underline{B}}) = \mathrm{e}^{\mathrm{Tr}\underline{\underline{A}}}.$$

24. If $\hat{P}_i|\psi\rangle = |\psi_i\rangle\langle\psi_i|\psi\rangle$, where the set $\{|\psi_i\rangle, i = 1, 2, \cdots, n\}$ forms a complete orthonormal basis in an n-dimensional vector space, show that
    (a) $\sum_{i=1}^{n} \hat{P}_i = \hat{1}$
    (b) $\hat{P}_i\hat{P}_j = \hat{P}_j\hat{P}_i = 0 \quad \text{for } i \neq j$
25. Find the matrix representation of the operator $\hat{A} = |\alpha\rangle\langle\beta|$ in a given discrete orthonormal basis $\{|\phi_i\rangle, i = 1, 2, \cdots\}$, where $|\alpha\rangle$ and $|\beta\rangle$ are two arbitrary kets in the Hilbert space.
26. If $|i\rangle$ and $|j\rangle$ are eigen kets of some Hermitian operator $\hat{A}$, under what condition $|i\rangle + |j\rangle$ is also an eigen ket of $\hat{A}$?
27. Consider the ket space spanned by the eigen kets $|\alpha_i\rangle$ (corresponding to eigen value a_i) of a Hermitian operator $\hat{A}$. There is no degeneracy.

(a) Prove that

$$\prod_i (\hat{A} - a_i) \text{ is the null operator}$$

(b) What is the significance of the operator

$$\prod_{i \neq j} \frac{(\hat{A} - a_i)}{(a_j - a_i)}$$

References

Arfken, G.: Mathematical Methods for Physicists. Academic Press, New York (1966)
Chattopadhyay, P.K.: Mathematical Physics. New Age International (P) Ltd., New Delhi (2006)

Chapter 3
Axiomatic Approach to Quantum Mechanics



Abstract This chapter develops modern quantum mechanics from its fundamental postulates, according to the concepts of linear vector space. It then develops quantization rules, time evolution of state vectors, coordinate space wave function, etc., in logical steps, without any assumptions or external inputs. Dirac delta function is introduced as a mathematical preliminary.

Keywords Axiomatic QM · Postulates of QM · Hilbert space · Measurement theory · Expansion postulate · Quantization · Commutation relations · Schrödinger eqution · Time evolution · Wave function · Equation of continuity · Dirac delta function

3.1 Linear Vector Spaces in Quantum Mechanics

We discussed in Chap. 1, that quantum mechanics was developed in a well defined fashion by Schrödinger, Heisenberg and others to explain observations involving 'microscopic' (atomic and subatomic) systems. Schrödinger's approach introduced the 'wave function', $\psi(\vec{r}, t)$, which is a mathematical function of the position $\vec{r}$ of the particle and time t. It provides all quantum mechanically admissible information about its motion in a potential $V(\vec{r}, t)$. This wave function is obtained by solving a differential equation called 'Schrödinger equation', Eq. (1.6), subject to specific boundary conditions. This is the governing equation of motion of this approach. Note that this equation is *linear and homogeneous* in $\psi(\vec{r}, t)$ and so ψ is normalizable. Also $\psi(\vec{r}, t)$ must be 'well behaved'; i.e. it must be square integrable and both ψ and its first derivatives must be continuous (the derivative may be discontinuous only at an infinite potential step). We saw in Chap. 2, that such a set of functions forms a specific linear vector space, called Hilbert space. An observable, represented in general by a differential operator, acts on this space. Such operators also satisfy eigen value equations. A measurement of this observable yields one of its eigen values.

In Heisenberg's *matrix mechanics* approach, the state of a system is represented (in a chosen basis) by a column matrix and all physically observable quantities are represented by Hermitian square matrices. In particular, energy of the system is repre-

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5_3

sented by the Hamiltonian matrix. A possible result of measurement of an observable is one of its eigen values. The matrix eigen value equation of the Hamiltonian ($\underline{\underline{H}}$) is the governing equation of motion for the system. The eigen values of $\underline{\underline{H}}$ are the possible results of a measurement of energy. The eigen column matrix corresponding to a particular eigen value E_i represents the state of the physical system corresponding to energy E_i. The set of all such column matrices forms a Hilbert space and Hermitian square matrices representing observables act on this space.

Both approaches yield the same result, and they are completely equivalent. Both were equally successful in explaining experimental observations. This shows that the Hilbert space in both the approaches is the same. An abstract Hilbert space ($\mathcal{H}$) is associated with the motion of a microscopic system. The governing equation of motion of the system in this $\mathcal{H}$ is the abstract Schrödinger equation(see below). Projection of this abstract equation on to the coordinate space and the matrix space are the coordinate space Schrödinger equationand the eigen value equation of the Hamiltonian matrix respectively.

Although both the wave mechanics and the matrix mechanics approaches were very successful in explaining experimental observations, the underlying mathematical basis became clear only later. It became clear that realizable states of a physical system belong to $\mathcal{H}$ and observables are operators acting on this space, taking one physical state to another. These ideas were later organized into a set of *fundamental postulates of quantum mechanics*, which were the basic assumptions and specified the rules for obtaining physical results. The postulates refer to abstract vectors in the abstract vector space $\mathcal{H}$, on which abstract operators act. We will see that coordinate or momentum representations of these abstract vectors correspond respectively to the coordinate or momentum space wave functions of the system. This procedure gives rise to the Schrödinger approach. Alternatively, we can have a matrix representation of $\mathcal{H}$, in a particular basis. This gives rise to the Heisenberg approach. We saw that vectors and operators in an abstract vector space can be represented by column and square matrices respectively, in this matrix space. Thus the approaches of Schrödinger and Heisenberg are equivalent. One can have other possibilities. For example, operators corresponding to physical quantities can be represented by *integral operators* and the differential Schrödinger equation is replaced by an *integral equation*, called Lippmann–Schwinger equation (Chap. 14). Thus all such approaches are equivalent.

In the next section, we discuss the fundamental postulates and see how they give rise to different approaches of quantum mechanics. For elegance, we adopt Dirac's bra-ket notations.

3.2 Fundamental Postulates of Quantum Mechanics

Generalizing the ideas of the previous section, several 'fundamental postulates' can be framed, which define the basic rules of quantum mechanics and how these rules can be associated with experimental measurements (Schiff 1955). In the following we will discuss these for a single particle system. A many-body system can be generalized from these.

First we will define the terminology to be used.

By the general name *physical system* we mean a particle or a group of particles (in general microscopic, i.e. atomic or subatomic). The motion of such a system is governed by the laws of quantum mechanics, which in turn are dictated by the *fundamental postulates*. Measurable physical quantities like energy of the 'system' and probability of finding the 'system' at a particular position and time are referred to as *observables*. The system can be in one of various possible '*states*'. Values of observables for the system depend on the *state of the system*.

In the following, we enumerate the fundamental axiomatic assumptions which guide the rules of quantum mechanics as the basic postulates:

- *Postulate 1:* A physical system and its state
  This postulate asserts that an abstract Hilbert space $\mathcal{H}$ can be associated with every physical system. A possible realizable state of the system is represented by a vector (say, $|\psi\rangle$) in this space and *vice versa*; i.e. every vector in $\mathcal{H}$ corresponds to a realizable state of the system.
- *Postulate 2:* An observable and the result of its measurement
  This postulate asserts that every observable (i.e. a physically measurable quantity), a, of the physical system is represented by a linear Hermitian operator, $\hat{A}$, acting on $\mathcal{H}$, and satisfies an eigen value equation

$$\hat{A}|\phi_i\rangle = a_i|\phi_i\rangle, \quad (i = 1, 2, \cdots). \tag{3.1}$$

  Since $\hat{A}$ is Hermitian, the set of all eigen kets $\{|\phi_i\rangle, i = 1, 2, \cdots\}$ forms a complete set. Hence this set spans $\mathcal{H}$. Eigen kets belonging to different eigen values are orthogonal (see Chap. 2, Sect. 2.3.1). They are assumed to be orthonormalized.
  Postulate 2 further asserts: The *only* possible result of a measurement of the observable (a) for the physical system will be *one of the eigen values* $\{a_i, \quad i = 1, 2, \cdots\}$ of $\hat{A}$.
  In Chap. 2, Sect. 2.3.1, we saw that the eigen values of a Hermitian operator are real numbers. Since the result of a measurement must be a real number, it is appropriate that observables be represented by Hermitian operators. Furthermore, eigen vectors (i.e. eigen kets) belonging to different eigen values are orthogonal (and assumed normalized)

$$\langle\phi_i|\phi_j\rangle = \delta_{ij} \quad (\text{for } a_i \neq a_j).$$

  If an eigen value appears more than once (called degenerate eigen value), corresponding eigen vectors are not automatically orthogonal, but are linearly independent and they can be orthonormalized (see Chap. 2, Sect. 2.3.1). Furthermore the set of all eigen vectors forms a complete set, and hence it satisfies the closure relation, Eq. (2.49)

$$\sum_i |\phi_i\rangle\langle\phi_i| = \hat{1}.$$

  Thus the set of all eigen vectors of $\hat{A}$ forms a complete orthonormal set (CONS).

– *Corollary 2a*: Expansion postulate
A state vector $|\psi\rangle$ (assumed normalized) belonging to $\mathcal{H}$, and hence representing a particular state of the system, can be expanded in the complete set of eigen kets $\{|\phi_i\rangle, \quad i = 1, 2, \cdots\}$ of the operator $\hat{A}$ representing an observable a

$$|\psi\rangle = \sum_i c_i |\phi_i\rangle, \tag{3.2}$$

where c_i is a constant (complex in general).
– *Corollary 2b*: Outcome of a measurement
If the result of a measurement of the observable a for this system yields the particular eigen value a_i, then Postulate 2 further asserts that the state of the system changes drastically ('the state collapses') to the eigen ket $|\phi_i\rangle$ of $\hat{A}$ belonging to the eigen value a_i, immediately after the 'measurement'. *Thus a 'measurement on the system' interferes with the system and changes the 'state' of the system.* This is in marked contrast with 'classical physics', where an external measurement does not change the state of the system.
If the initial state of the system is represented by $|\psi\rangle$, which is *not* an eigen ket of $\hat{A}$, the probability of getting a particular eigen value a_j is $|c_j|^2$, where c_j is the expansion coefficient of $|\psi\rangle$ in the complete set of eigen kets $\{|\phi_i\rangle, i = 1, 2, \cdots\}$ of $\hat{A}$ [see Eq. (3.2)]. Since the system then *collapses into the eigen vector* $|\phi_j\rangle$*,any subsequent measurement of the same observable will yield the same eigen value* a_j *each time, with certainty*. If the system is initially in the eigen ket $|\phi_k\rangle$, then a measurement of a will yield the value a_k with certainty.

Consequence of Corollary 2b: If the initial state of the system $|\psi\rangle$ is *not* an eigen vector of $\hat{A}$, we can *only give the probability* for the result of a measurement of a. This brings in a *probabilistic interpretation* of $|\psi\rangle$, which is further clarified in Corollary 2c (see below).
This postulate makes the quantum mechanics drastically different from the classical mechanics. It says that for microscopic systems, a measurement is *not independent of the system*, rather it changes the state of the system. We can understand this from the following crude analogy. Consider a one-gram ball moving with a speed v in the x-direction. We can measure its speed by 'seeing' positions of the ball at two different times. This does not change the speed of the ball in our everyday experience. But how do we 'see' the ball? By light reflected from the ball and it does not change its speed. Now imagine that the mass of the ball in reduced to that of an electron. To measure speed, we have to 'see' this tiny ball at two different times. But as light is reflected from the tiny ball, the ball recoils (hence its momentum changes), since the momentum of the light photon ($\frac{h\nu}{c}$) is now *comparable* with the momentum of the tiny ball. Thus as we 'measure' the position or momentum of a microscopic system, the latter is 'disturbed' and its state changes. Incidentally, this also shows that the 'speed' of a quantum particle is not a measurable quantity.

– *Corollary 2c:* Expectation value.
If the system is in a state $|\psi\rangle$, which is *not* an eigen state of $\hat{A}$, what will be the *average* outcome or *the expectation value* of a measurement of a? This is given by

$$\langle \hat{A} \rangle = \frac{\langle \psi | \hat{A} | \psi \rangle}{\langle \psi | \psi \rangle}. \tag{3.3}$$

(The denominator is needed to *normalize* the state $|\psi\rangle$, if it is not normalized.) What is the physical meaning of this 'expectation value'? By this, we mean $\langle \hat{A} \rangle$ is the 'average result' (i.e. the 'expected value') of a large number of measurements of a on the system in the 'state' $|\psi\rangle$. However, this appears to be in conflict with Corollary 2b (if after getting an eigen value a_j, the system 'collapses' to $|\phi_j\rangle$, and *all* subsequent measurements on *this collapsed system* will yield the value a_j *only*, then the 'average' would eventually be a_j). This is resolved by giving a 'statistical meaning' to $|\psi\rangle$. When we talk about the 'expectation value', we mean repeated measurements on the system in the **same state** $|\psi\rangle$ before each measurement. Thus we imagine a very large number of *identical copies* of the system in the same state (so-called 'identically prepared systems') and the observable a is measured *independently* in each one of these 'identical copies'. After a *single* measurement on a particular (say, k-th) copy, the result may be a_k with a probability $P(a_k) = |c_k|^2$, where c_k is given by Eq. (3.2), according to Corollary 2b. Hence the average value will be

$$a_{av} = \frac{\sum_k a_k |c_k|^2}{\sum_k |c_k|^2} \tag{3.4}$$

(If $|\psi\rangle$ and all $|\phi_i\rangle$ are normalized, the denominator is unity.) Substituting Eqs. (3.1) and (3.2) in Eq. (3.3) and and comparing with Eq. (3.4), it is seen that $a_{av} = \langle \hat{A} \rangle$. Hence the expectation value is also called the 'ensemble average' of measurement of a for the system in the state $|\psi\rangle$.

- *Postulate 3:* Quantization postulate
This postulate asserts: The pair of operators representing a classical generalized variable, $\{q_i\}$ and its canonically conjugate momentum $\{p_i\}$, do not commute and satisfy commutation relations (since they do not commute, they cannot be numbers, but are operators denoted by a hat ($\hat{q}_i$ or $\hat{p}_j$) above the symbol):

$$\begin{aligned} \left[\hat{q}_i, \hat{q}_j\right] = 0 = \left[\hat{p}_i, \hat{p}_j\right], \\ \left[\hat{q}_i, \hat{p}_j\right] = i\hbar\delta_{ij}, \end{aligned} \tag{3.5}$$

where $\hbar = \frac{h}{2\pi}$, h being Planck's constant ($h = 6.622 \times 10^{-27}$ erg.sec). As a special case, Cartesian components of position ($\hat{\vec{r}}$) and momentum ($\hat{\vec{p}}$) operators satisfy

$$\left[\hat{x}_i, \hat{p}_j\right] = i\hbar\delta_{ij}, \tag{3.6}$$

where x_i and p_i are the i-th Cartesian component ($i = 1, 2, 3$ corresponding to x, y, z) of $\vec{r}$ and $\vec{p}$ respectively. This is called *first quantization* (to distinguish it from *second quantization* of the fields).
In *coordinate representation*, where the fundamental variable is the position vector, $\vec{r}$ (see below and Chap. 4, Sect. 4.1), Eq. (3.6) is satisfied by replacing $\vec{p}$ by the differential operator

$$\hat{\vec{p}} = -i\hbar\vec{\nabla}, \tag{3.7}$$

so that components of $\vec{r}$ and $\vec{p}$ satisfy the commutation relations (3.6). This can be verified by applying Eq. (3.6) for the x-component on a general state vector $\psi(\vec{r})$ in the coordinate representation:

$$\begin{aligned}[\hat{x}, \hat{p}_x]\psi(\vec{r}) &= \left(x(-i\hbar\frac{\partial}{\partial x}) - (-i\hbar\frac{\partial}{\partial x})x\right)\psi(x, y, z) \\ &= -i\hbar\left(x\frac{\partial\psi}{\partial x} - (x\frac{\partial\psi}{\partial x} + \psi)\right) = i\hbar\psi.\end{aligned}$$

This is true for any ψ. Hence we have

$$[\hat{x}, \hat{p}_x] = i\hbar,$$

proving x-component of Eq. (3.6). We can prove similarly for other components.

Since $\hat{x}_i$ and $\hat{p}_i$ do not commute [Eq. (3.6)], they cannot have simultaneously eigen vectors (see Chap. 2, Sect. 2.3.1) and according to Postulate 2, they cannot simultaneously be specified precisely. This result leads to the famous *Heisenberg's position-momentum uncertainty relation* (see Chap. 5)

$$\Delta x\ \Delta p_x \geq \frac{\hbar}{2},$$

where Δx and Δp_x are measures of uncertainty (root-mean-square deviation from the mean) in the *simultaneous average measurements* of x and p_x respectively.
In a similar fashion, if the fundamental variable is the momentum $\vec{p}$ (see Chap. 4, Sect. 4.2), position vector has to be represented by a different differential operator acting on $\vec{p}$

$$\hat{\vec{r}} = i\hbar\vec{\nabla}_{\vec{p}},$$

(where $\vec{\nabla}_{\vec{p}}$ is the gradient operator corresponding to variable $\vec{p}$) such that the commutations (3.6) remain valid (which can be verified as above).

Special quantization for the dynamical variable 'energy':
Energy is represented by

$$\hat{E} \rightarrow i\hbar \frac{\partial}{\partial t}. \tag{3.8}$$

This is borrowed from classical mechanics in which the Hamiltonian corresponds to energy of the system and is the generator of infinitesimal time evolution (see Ref. Goldstein 1950). Note that as in classical mechanics, in non-relativistic quantum mechanics also, time (t) is treated as an external parameter and not a dynamical variable, and so a commutation relation between $\hat{E}$ and t (*viz*, $[t, \hat{E}] = -i\hbar$), analogous to Eqs. (3.5) and (3.6)) is not meaningful in the present context (for more on this see Chap. 2 of Ref. Sakurai 2000). However, we will see in Chap. 5 that Eq. (3.8) leads to the time-energy uncertainty relation

$$\Delta t \, \Delta E \geq \frac{\hbar}{2},$$

which has been verified experimentally.
A corollary follows Postulate 3:

– *Corollary 3a:* Prescription to obtain an operator representing a dynamical variable
The quantum operator corresponding to an observable is obtained from its classical expression, replacing the generalized coordinates and their canonically conjugate momenta by their corresponding operators satisfying the commutation relations Eq. (3.5).
This leads to different 'representations', depending on the choice of generalized coordinates and momenta. Choice of positions ($\hat{\vec{r}}$) as the generalized variable and its conjugate momenta ($\hat{\vec{p}} = -i\hbar\vec{\nabla}$) gives rise to the commonly used 'coordinate representation'. On the other hand, choosing momentum ($\hat{\vec{p}}$) as the generalized variable and its conjugate variable $\hat{\vec{r}} = i\hbar\hat{\vec{\nabla}}_{\vec{p}}$ gives the 'momentum representation'.
Thus in coordinate representation, which is commonly used, energy of the system is represented by the Hamiltonian operator, obtained from its classical expression, with $\hat{\vec{r}}$ as the generalized variable and $\hat{\vec{p}} = -i\hbar\hat{\vec{\nabla}}$ as the conjugate variable,

$$H = \frac{p^2}{2m} + V(\vec{r}, t) \quad \text{(classical)}, \tag{3.9}$$

$$\hat{H} = -\frac{\hbar^2}{2m}\hat{\nabla}^2 + V(\hat{\vec{r}}, t) \quad \text{(quantum mechanical)}. \tag{3.10}$$

- *Postulate 4:* Prescription for obtaining the state of a physical system
From the general observation (leading to an important principle in physics) that nature always prefers any system, left to itself, to be in its lowest possible energy state, it is apparent that the lowest energy state of any system has a special impor-

tance in physics. Thus the **energy** of the system is the main decider of the state of the system. In quantum mechanics energy is represented by the Hamiltonian operator, $\hat{H}$
Hence this postulate asserts that the state $|\psi\rangle$ of the system is given as the 'energy eigen state', i.e. *eigen vector* of $\hat{H}$ corresponding to eigen value E

$$\hat{H}|\psi\rangle = E|\psi\rangle. \tag{3.11}$$

In coordinate representation, this is given by, using Eq. (3.10),

$$\left[-\frac{\hbar^2}{2m}\hat{\nabla}^2 + V(\hat{\vec{r}}, t_0)\right]\psi(\vec{r}) = E\psi(\vec{r}), \tag{3.12}$$

where $\psi(\vec{r})$ is the 'wave function' (coordinate projection of $|\psi\rangle$, see Chap. 4, Sect. 4.1) at *some fixed time* t_0. Equation (3.12) is the well known time-independent Schrödinger equation in coordinate representation, while Eq. (3.11) is the Schrödinger equation in the Hilbert space.

- *Corollary 4a:* Time evolution of the state vector, $|\psi\rangle$
  If the state of a system is given by $|\psi(t_0)\rangle$ at some time t_0, how does it evolve in time? The answer is given in the following treatment. We expect that there is an operator $\hat{U}(t, t_0)$ (called *time evolution operator*) in $\mathcal{H}$, which acting on $|\psi(t_0)\rangle$ takes it to $|\psi(t)\rangle$

$$|\psi(t)\rangle = \hat{U}(t, t_0)|\psi(t_0)\rangle. \tag{3.13}$$

  Replacing t_0 by t, we have

$$\hat{U}(t, t) = \hat{1}. \tag{3.14}$$

  As the total probability for finding the system should remain the same, we expect that if $|\psi(t_0)\rangle$ is normalized, $|\psi(t)\rangle$ will also be normalized. Hence,

$$\langle\psi(t)|\psi(t)\rangle = \langle\psi(t_0)|\hat{U}^\dagger(t, t_0)\hat{U}(t, t_0)|\psi(t_0)\rangle = \langle\psi(t_0)|\psi(t_0)\rangle.$$

  Since this is true for all $|\psi(t_0)\rangle$, we have

$$\hat{U}^\dagger(t, t_0)\hat{U}(t, t_0) = \hat{1}. \tag{3.15}$$

  Thus $\hat{U}(t, t_0)$ is unitary.
  We also want $\hat{U}(t, t_0)$ to have the property

$$\hat{U}(t_2, t_0) = \hat{U}(t_2, t_1)\hat{U}(t_1, t_0), \quad (t_2 > t_1 > t_0), \tag{3.16}$$

  so that the time evolution from t_0 to t_2 can be in two steps: first from t_0 to t_1, then again from t_1 to t_2. Furthermore, time is a continuous variable. For an infinitesimal time evolution through $\mathrm{d}t$, we have intuitively

$$\lim_{\mathrm{d}t\to 0} \hat{U}(t+\mathrm{d}t, t) = \hat{1}. \tag{3.17}$$

Expanding $\hat{U}(t+\mathrm{d}t, t)$ in a power series of $\mathrm{d}t$ and disregarding terms containing $(\mathrm{d}t)^2$ and higher powers for an infinitesimal $\mathrm{d}t$, we can write

$$\hat{U}(t+\mathrm{d}t, t) = \hat{1} - i\hat{\Omega}\mathrm{d}t,$$

where we write the coefficient of $\mathrm{d}t$ as $-i\hat{\Omega}$, for reasons to be clear in the following. We can easily verify that it will satisfy properties (3.14) and (3.16)–(3.17). This choice shows that the operator $\hat{\Omega}$ is the generator of an infinitesimal time evolution. For the property (3.15), $\hat{\Omega}$ has to be Hermitian

$$\hat{\Omega}^\dagger = \hat{\Omega}.$$

The operator $\hat{\Omega}$ has the dimension of frequency or inverse time (since $i\hat{\Omega}\mathrm{d}t$ has to be dimensionless, like $\hat{1}$). In classical mechanics, the Hamiltonian is the generator of infinitesimal time evolution (see Ref. Goldstein 1950). We therefore relate $\hat{\Omega}$ to the Hamiltonian $\hat{H}$ of the physical system divided by $\hbar$ (to give it the appropriate dimension), $\hat{\Omega} = \frac{\hat{H}}{\hbar}$. Thus

$$\hat{U}(t+\mathrm{d}t, t) = \hat{1} - i\frac{\hat{H}}{\hbar}\mathrm{d}t. \tag{3.18}$$

Putting $t_1 = t$ and $t_2 = t + \mathrm{d}t$ in Eq. (3.16) and using the last equation, we have

$$\hat{U}(t+\mathrm{d}t, t_0) = \hat{U}(t+\mathrm{d}t, t)\hat{U}(t, t_0) = \left(\hat{1} - i\frac{\hat{H}}{\hbar}\mathrm{d}t\right)\hat{U}(t, t_0).$$

Hence

$$\hat{U}(t+\mathrm{d}t, t_0) - \hat{U}(t, t_0) = -i\frac{\hat{H}}{\hbar}\mathrm{d}t\hat{U}(t, t_0).$$

Dividing by $-\frac{i}{\hbar}\mathrm{d}t$ and taking the limit $\mathrm{d}t \to 0$, we have

$$i\hbar\frac{\partial}{\partial t}\hat{U}(t, t_0) = \hat{H}\hat{U}(t, t_0). \tag{3.19}$$

This is the equation satisfied by the operator for time evolution from time t_0 to time t. Note that the difference $(t - t_0)$ need not be infinitesimal.

If Hamiltonian, $\hat{H}$, is independent of time, we can obtain the time evolution operator for a finite time interval, $(t - t_0)$, by first dividing the time interval in N steps of $(\frac{t-t_0}{N})$ and letting $N \to \infty$, so that the time interval $(\frac{t-t_0}{N}) = \Delta t$ becomes infinitesimal. Then applying the infinitesimal time evolution operator, Eq. (3.18), successively at each step (note that it is independent of the initial time) and using Eq. (3.16), the total time evolution operator is the product of an infinite number of infinitesimal time evolution operators

$$\begin{aligned}\hat{U}(t, t_0) &= \lim_{N\to\infty} \left[\hat{U}(t_0 + \Delta t, t_0)\right]^N \\ &= \lim_{N\to\infty} \left[\hat{1} - \frac{i\hat{H}\Delta t}{\hbar}\right]^N \\ &= \lim_{N\to\infty} \left[\hat{1} - \frac{i\hat{H}}{\hbar}\frac{(t - t_0)}{N}\right]^N \\ &= \exp\left(-\frac{i\hat{H}(t - t_0)}{\hbar}\right) \end{aligned} \tag{3.20}$$

(using $e^x = \lim_{N\to\infty}\left[1 + \frac{x}{N}\right]^N$, which is true also for operators).
Applying Eq. (3.19) on $|\psi(t_0)\rangle$ and using Eq. (3.13) we get

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle. \tag{3.21}$$

This is the Hilbert space version of the celebrated Schrödinger equation for the time evolution for the state of a physical system, popularly known as time-dependent Schrödinger equation. In the above treatment, we apparently 'derived' the Schrödinger equation. However, note that this equation, like other 'laws of physics' is a *generalization of experience* and cannot be 'derived'. Equation (3.11) is the generalization of observations. The main basis of the treatment of time evolution is Eq. (3.18), which borrows the classical result that the Hamiltonian is the generator of infinitesimal time evolution.

3.3 Coordinate Space Wave Function: Interpretation

'Position' of the particle has to be specified in the 'position representation', which is popularly called *the coordinate space representation*. Hence we consider the position operator $\hat{\vec{r}}$ in coordinate space. It is Hermitian, as it is an observable and satisfies an eigen value equation with eigen value $\vec{r}\,'$ (a continuous variable) corresponding to the eigen ket $|\vec{r}\,'\rangle$ (the eigen ket is characterized, i.e. named by its eigen value)

$$\hat{\vec{r}}\,|\vec{r}\,'\rangle = \vec{r}\,'|\vec{r}\,'\rangle.$$

In view of continuity of the eigen value, the orthonormality of these eigen kets is given by (see Chap. 4, Sect. 4.1)

$$\langle \vec{r}\,' | \vec{r}\,'' \rangle = \delta^{(3)}(\vec{r}\,' - \vec{r}\,''), \tag{3.22}$$

where $\delta^{(3)}(\vec{r}\,' - \vec{r}\,'')$ is Dirac's delta function (see Sect. 3.4). The completeness of $\{|\vec{r}\,'\rangle\}$ is given by

$$\int |\vec{r}\,'\rangle \, \mathrm{d}^3 r' \, \langle \vec{r}\,' | = \hat{1}, \tag{3.23}$$

where the integral is over all space. We will follow the convention: a definite integral without limits specified, means the integral is over all space of the variable of integration. According to the expansion postulate, we can expand the state of the system $|\psi(t)\rangle$ (including time dependence) in the complete set of position eigen kets. This can be done by inserting Eq. (3.23) in front of the state ket

$$|\psi(t)\rangle = \int |\vec{r}\,\rangle \, \mathrm{d}^3 r \, \langle \vec{r} \,|\, \psi(t)\rangle \equiv \int |\vec{r}\,\rangle \, \mathrm{d}^3 r \, \psi(\vec{r}, t),$$
$$\text{where} \quad \psi(\vec{r}, t) = \langle \vec{r} \,|\, \psi(t)\rangle. \tag{3.24}$$

Hence $\psi(\vec{r}, t)$ is the expansion coefficient. It is called the *coordinate space wave function* and is a function of position $\vec{r}$ and time t. It is also the coordinate space projection of the abstract state vector $|\psi(t)\rangle$. According to Corollary 2b of this postulate, $|\psi(\vec{r}, t)|^2 \mathrm{d}^3 r$ is the probability that a position measurement on the system (in the state $|\psi(t)\rangle$) will yield a value $\vec{r}$ within a small volume element $\mathrm{d}^3 r$. In other words, the probability of finding the particle within a small volume element $\mathrm{d}^3 r$ around the position $\vec{r}$ at time t is $|\psi(\vec{r}, t)|^2 \mathrm{d}^3 r$. The quantity $|\psi(\vec{r}, t)|^2$ is called the position probability density. This is the *probabilistic interpretation of the wave function*. Thus $|\psi(\vec{r}, t)|^2$ must be finite and the total probability of finding the particle *anywhere in space* at any time t should be unity (i.e. a certainty)

$$\int |\psi(\vec{r}, t)|^2 \mathrm{d}^3 r = 1, \tag{3.25}$$

where the integral is over all space. Equation (3.21) satisfied by the state ket $|\psi(t)\rangle$ is linear and homogeneous in the state ket. Hence it is normalizable by a multiplying constant. Consequently Eq. (3.24) shows that the wave function $\psi(\vec{r}, t)$ is also normalizable. It is normalized to one according to Eq. (3.25). Equation (3.24) together with Eqs. (3.22), (3.37) (property of the δ-function, see below) and (3.25) show that $|\psi(t)\rangle$ is normalized according to

$$\langle \psi(t) | \psi(t) \rangle = 1.$$

Bound and unbound states

A system under consideration may be localized or not.

A '*localized system*' is one which has zero probability of finding it at an infinite distance from a localized (i.e. finite) region of space, so that $\psi(\vec{r}, t) \to 0$ for $r \to \infty$. In this case the integral on the left of Eq. (3.25) is finite. It is possible in two ways:

1. For all t, the wave function is localized in space, i.e. $\psi(\vec{r}, t) \to 0$ for $r \to \infty$ for all t. As an example a particle confined in a deep well, fixed in space, cannot escape to infinity. Such a system is called a *bound system*. If the potential $V(\vec{r}, t)$ is independent of t, then time dependence of ψ factors out and $|\psi(\vec{r}, t)|^2$ becomes independent of t (see Chap. 7).
2. Another possible localized system may be a system that is moving in time, but at any given time t, it is localized in space and Eq. (3.25) can be satisfied at any particular t. A moving particle is an example of this type. Corresponding wave function is called a '*traveling wave packet*'. This can also happen to a linear combination of energy eigen functions, as can be seen for the harmonic oscillator (see Chap. 9, Sect. 9.4).

On the other hand, for a system which is *not localized*, $\psi(\vec{r}, t)$ does not vanish even as $r \to \infty$, at any time t and the integral in Eq. (3.25) diverges. Hence this condition cannot be satisfied. But Eq. (3.21), which is linear and homogeneous in $|\psi(t)\rangle$, still holds and hence $\psi(\vec{r}, t)$ can still be normalized by other conditions, as we will discuss in Sect. 3.5. Scattering of a particle coming from infinity, interacting with another particle at the origin and moving on to infinity (in general in a different direction) is an example of this type (we will discuss this in Chap. 14).

This interpretation of $\psi(\vec{r}, t)$ also suggests that the *probability density* $\rho(\vec{r}, t)$ (which is the probability per unit volume) of finding the particle at a position $\vec{r}$ at time t is

$$\rho(\vec{r}, t) = |\psi(\vec{r}, t)|^2. \tag{3.26}$$

We next obtain a differential equation satisfied by the wave function. To do this, we take inner product of Eq. (3.21) with $\langle \vec{r}\,|$ (note that this is independent of t), insert Eq. (3.23) and use Eq. (3.24) to get

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, t) = \langle \vec{r}\,|\hat{H}(t) \int |\vec{r}\,'\rangle \mathrm{d}^3\vec{r}\,' \langle \vec{r}\,'|\psi(t)\rangle,$$

where a possible time dependence of the Hamiltonian [through time dependence of the potential $V(\vec{r}, t)$] has been explicitly shown. Using

$$\langle \vec{r}\,|\,\hat{H}(t)\,|\,\vec{r}\,'\rangle = H(\vec{r}, t)\,\delta(\vec{r} - \vec{r}\,')$$

for a *local Hamiltonian*, we have using Eq. (3.10)

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = H(\vec{r},t)\psi(\vec{r},t) = \Big[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r},t)\Big]\psi(\vec{r},t). \tag{3.27}$$

This is known as the *coordinate space time-dependent Schrödinger equation.*

Equation of continuity and probability current density

Differentiating Eq. (3.26) partially with respect to time and substituting Eq. (3.27) and its complex conjugate equation, we have (assuming $V(\vec{r},t)$ to be real)

$$\begin{aligned}\frac{\partial\rho(\vec{r},t)}{\partial t} &= \psi^*(\vec{r},t)\frac{\partial\psi(\vec{r},t)}{\partial t} + \frac{\partial\psi^*(\vec{r},t)}{\partial t}\psi(\vec{r},t)\\ &= \frac{i\hbar}{2m}\Big[\psi^*(\vec{r},t)\nabla^2\psi(\vec{r},t) - \Big(\nabla^2\psi^*(\vec{r},t)\Big)\psi(\vec{r},t)\Big]\\ &= -\frac{\hbar}{2im}\vec{\nabla}\cdot\Big[\psi^*(\vec{r},t)\vec{\nabla}\psi(\vec{r},t) - \Big(\vec{\nabla}\psi^*(\vec{r},t)\Big)\psi(\vec{r},t)\Big].\end{aligned} \tag{3.28}$$

Defining the *probability current density* as

$$\vec{j}(\vec{r},t) = \frac{\hbar}{2im}\Big[\psi^*(\vec{r},t)\vec{\nabla}\psi(\vec{r},t) - \Big(\vec{\nabla}\psi^*(\vec{r},t)\Big)\psi(\vec{r},t)\Big], \tag{3.29}$$

we obtain the equation of continuity for probability density

$$\frac{\partial\rho(\vec{r},t)}{\partial t} + \vec{\nabla}\cdot\vec{j}(\vec{r},t) = 0. \tag{3.30}$$

This is the familiar equation of continuity of an incompressible fluid of density $\rho(\vec{r},t)$ and current density $\vec{j}(\vec{r},t)$ when there are no sources or sinks of that fluid. This represents conservation of the fluid, as the rate of decrease in the amount of fluid in a small volume element d^3r about $\vec{r}$ must be the same as the rate of outflow of the fluid from that volume element, given by the divergence of the flux. Thus we can interpret $\vec{j}(\vec{r},t)$ given by Eq. (3.29) as the probability current density and Eq. (3.30) represents the conservation of quantum mechanical probability.

Flux or current density (defined as the quantity of fluid passing unit area perpendicular to the instantaneous direction of flow per unit time) of a classical fluid is given by $\vec{j}_{cl} = \rho\,\vec{v}_{cl} = \frac{1}{m}\rho\,\vec{p}_{cl}$, where $\vec{v}_{cl}$ and $\vec{p}_{cl}$ are the classical velocity and momentum respectively. Equation (3.29) can be written as $\vec{j}(\vec{r},t)$ = Real part of $\frac{1}{m}\Big[\psi^*(\vec{r},t)\Big(-i\hbar\vec{\nabla}\psi(\vec{r},t)\Big)\Big]$. With $\rho(\vec{r},t) = \psi^*(\vec{r},t)\psi(\vec{r},t)$, we see that $\vec{j}(\vec{r},t)$ given by Equation (3.29) represents the quantum equivalent of classical probability flux (i.e. probability current density). However, we note that this quantity is *not quantum mechanically measurable*, since a velocity measurement involves simultaneous knowledge of both position and momentum of the particle, which is not permitted by the uncertainty principle (as according to Postulate 3 operators for position and

momentum do not commute). It is a useful concept when the flux is independent of $\vec{r}$ (or depends only slightly on $\vec{r}$), as is used in flux normalization of a beam of particles (see Sect. 3.5).

Equation (3.30) represents conservation of probability for a bound state or a wave packet. It can be seen by integrating it over all space

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \rho(\vec{r},t)\,\mathrm{d}^3r = -\int \vec{\nabla}\cdot\vec{j}(\vec{r},t)\,\mathrm{d}^3r = -\oint \vec{j}(\vec{r},t)\cdot d\vec{A},$$

where we used Gauss' theorem in the second equality and $d\vec{A}$ is a surface element of the surface enclosing the entire volume of integration. As before the integral is over all space and $\oint$ is over the entire surface at infinity enclosing the entire space. For a bound state or a finite wave packet, $|\psi|$ vanishes at infinity. Hence the surface integral vanishes and we have

$$\int \rho(\vec{r},t)\,\mathrm{d}^3r = \int |\psi(\vec{r},t)|^2\mathrm{d}^3r = \text{constant} \quad \text{(independent of } t\text{)}.$$

This shows that, if the wave function is normalized according to Eq. (3.25) at a particular time t, it will remain normalized at all times, as the volume integral is independent of time. Note that this is true if the potential is real [see derivation of Eq. (3.28)]. An imaginary part in the potential leads to a source or sink of probability in the equation of continuity. A negative imaginary part of potential leads to decrease in the total probability with time, corresponding to absorption (see Problems, Sect. 3.6).

3.4 Mathematical Preliminary: Dirac Delta Function

In this section, we will discuss a special function, called *Dirac's delta function*, introduced by Dirac, which is very useful in continuous eigen value problems. This is a singular function, but can be represented as the limit of several analytic functions.

In one dimension the delta function is defined by

$$\delta(x) = 0 \ \text{ for } x \neq 0, \quad \text{such that } \int \delta(x)\mathrm{d}x = 1. \tag{3.31}$$

The domain of integration must include the point $x = 0$. Clearly $\delta(0)$ diverges, so that the integral is finite (equal to one). One also has

$$\int f(x)\delta(x)\mathrm{d}x = f(0), \tag{3.32}$$

where again the domain of integration includes the origin and $f(x)$ is continuous at $x = 0$. Equation (3.32) can be used as an alternative definition of δ-function, replacing the integral in the defining Eq. (3.31). We can visualize the δ-function as one which is zero at all points, except $x = 0$, where it is infinity, such that the integral

from $x = -\infty$ to $x - +\infty$, i.e. the area under the curve, is one. It is not analytic at $x = 0$. From Eqs. (3.31), (3.32) follows, since $f(x)$ is continuous at $x = 0$ and hence can be expanded in a Taylor series about $x = 0$ followed by the use of the integral in Eq. (3.31). Alternately setting $f(x) = 1$ in Eq. (3.32), we get Eq. (3.31).

The δ-function can be represented as limits of several analytic continuous or piecewise continuous functions:

1. Limit of an analytic function:

$$\delta(x) = \lim_{a\to\infty} \frac{\sin ax}{\pi x}. \tag{3.33}$$

The function at $x = 0$ has the value $\frac{a}{\pi}$, which goes to infinity in the limit $a \to \infty$. The function is symmetric about the origin, is oscillatory with period $\frac{2\pi}{a}$ and gradually decreasing amplitude. The period rapidly decreases to an infinitesimal value as $a \to \infty$, allowing each integral over a complete cycle to vanish. Then for $x \neq 0$ the function becomes negligible compared to its value $\frac{a}{\pi}$ at $x = 0$ in the limit $a \to \infty$. We also have

$$\begin{aligned}\int \delta(x)\mathrm{d}x &= \lim_{a\to\infty} \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin y}{y} dy, \quad \text{(substituting } y = ax\text{)} \\ &= \lim_{a\to\infty} \frac{1}{\pi}\, \pi \\ &= 1.\end{aligned}$$

Thus the function defined in Eq. (3.33) satisfies the definition of the δ-function and is a representation of it. Note that for a large but finite value of a, this function is analytic.

2. As an integral

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{ikx}\mathrm{d}k. \tag{3.34}$$

We can easily see that

$$\begin{aligned}\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{ikx}\mathrm{d}k &= \lim_{a\to\infty} \frac{1}{2\pi} \int_{-a}^{a} \mathrm{e}^{ikx}\mathrm{d}k \\ &= \lim_{a\to\infty} \frac{1}{\pi}\left(\frac{\mathrm{e}^{iax} - \mathrm{e}^{-iax}}{2ix}\right) = \lim_{a\to\infty} \frac{\sin ax}{\pi x} = \delta(x),\end{aligned}$$

by Eq. (3.33). Note that the sign of argument of the exponential can be either positive or negative, in agreement with the δ-function being an even function (see properties below).

There are several other representations of the delta function as limiting cases (Arfken 1966). The representation (3.34) is particularly useful in δ-function normalization in quantum mechanics (see the following section and Chap. 14).

Properties:
Some important properties of the δ-function are (for details, see Ref. Arfken 1966)

1. It is seen from definition (3.31) and Eqs. (3.32)-(3.34) that the δ-function is symmetric: $\delta(-x) = \delta(x)$.
2. Since $\delta(x) = 0$ for $x \neq 0$, $x\delta(x) = 0$.
3. Obviously, $\delta(bx)$ is proportional to $\delta(x)$. Equation (3.31) shows that $\delta(bx) = \frac{1}{b}\delta(x) \quad (b > 0)$.
4. We can prove by partial integration, using Eq. (3.31) [or better using Eq. (3.32)] that $x\delta'(x) = -\delta(x)$.
5. From the last relation, we see that the derivative of δ-function with respect to its argument is an odd function $\delta'(-x) = -\delta'(x)$.
6. Since the argument of $\delta(x^2 - b^2)$ vanishes for $x = \pm b$, it will be proportional to a sum of $\delta(x - b)$ and $\delta(x + b)$. The proportionality constant can be found from the def. (3.31) and we have $\delta(x^2 - b^2) = \frac{1}{2b}\Big(\delta(x - b) + \delta(x + b)\Big) \quad (b > 0)$.

Three-dimensional δ-function
In Cartesian system we can write

$$\delta^{(3)}(\vec{r}\,) = \delta(x)\delta(y)\delta(z). \tag{3.35}$$

Its properties are obtained intuitively from those of the one-dimensional δ-function. The general definition, analogous to Eq. (3.31) is

$$\delta^{(3)}(\vec{r}\,) = 0 \text{ for } r \neq 0 \text{ and } \int \delta^{(3)}(\vec{r}\,)\,\mathrm{d}^3r = \int \delta^{(3)}(\vec{r}\,)\,r^2\mathrm{d}r\,\sin\theta\mathrm{d}\theta\mathrm{d}\phi = 1. \tag{3.36}$$

In this, the integral is a three-dimensional integral over all space, i.e. 0 to ∞ for r, 0 to π for θ and 0 to 2π for ϕ. An alternative definition, analogous to Eq. (3.32) is

$$\int f(\vec{r}\,)\delta^{(3)}(\vec{r} - \vec{r}\,')\,\mathrm{d}^3r = \int f(\vec{r}\,)\delta^{(3)}(\vec{r} - \vec{r}\,')\,r^2\mathrm{d}r\,\sin\theta\mathrm{d}\theta\mathrm{d}\phi = f(\vec{r}\,'), \tag{3.37}$$

where $f(\vec{r})$ is continuous around $\vec{r} = \vec{r}\,'$. Using Eq. (3.34) for x, y and z in Eq. (3.35), we have for three-dimensional δ-function

$$\begin{aligned}\delta^{(3)}(\vec{r}\,) &= \frac{1}{(2\pi)^3}\int \mathrm{e}^{i\vec{k}.\vec{r}}\mathrm{d}^3k,\\ \delta^{(3)}(\vec{k}) &= \frac{1}{(2\pi)^3}\int \mathrm{e}^{-i\vec{k}.\vec{r}}\mathrm{d}^3r.\end{aligned} \tag{3.38}$$

The last relation is obtained by interchanging $\vec{r}$ and $\vec{k}$ in the previous one and then replacing $\vec{r}$ by $-\vec{r}$. This is done for convenience in going from coordinate to momentum representations and *vice versa* (see Chap. 4). Note that the sign of argument of the exponential in the above equations, as well as Eq. (3.34) can be either positive or negative, as the δ-functions are symmetric in their arguments.

3.5 Normalization

For a single particle bound in space by a confining potential (which we take to be real), the wave function is normalized at time t according to Eq. (3.25), where the integral is over all space. The discussion at the end of Sect. 3.3 tells us that a normalized wave function will remain normalized at all times for real potentials. This is true for both types of localized systems discussed in Sect. 3.3.

However, there may be non-localized physical situations in which the wave function does not vanish at infinity, but has a finite value and the integral over all space diverges. Such a situation arises for a traveling plane wave

$$\psi_{TW}(\vec{r}, t) = A\mathrm{e}^{i(\vec{k}\cdot\vec{r}-\omega t)}, \tag{3.39}$$

where A is an arbitrary constant. Application of the momentum operator $-i\hbar\vec{\nabla}$ gives the eigen value $\hbar\vec{k}$ and using Eq. (3.29) we get $\vec{j}(\vec{r}, t) = \frac{\hbar\vec{k}}{m}|A|^2 \equiv \vec{v}_{in}|A|^2$. Hence ψ_{TW} represents a plane wave of constant amplitude A traveling in the direction $\vec{k}$ with a flux of constant magnitude $\frac{\hbar k}{m}|A|^2$. Note that $\mathrm{e}^{i(\vec{k}\cdot\vec{r}-\omega t)}$ corresponds to a *definite momentum* $\hbar\vec{k}$, completely unspecified position (in agreement with uncertainty principle) and probability density 1. Hence its flux[1] can be specified as $\frac{\hbar\vec{k}}{m}$, which is the classical velocity $\vec{v}_{cl}$ (we call it 'the incident flux' $\vec{v}_{in}$). A similar situation will arise in a scattering problem, in which a particle comes from infinity, interacts with a localized potential field and then flies away to infinity. The time-independent wave function of the particle will not vanish at infinity and Eq. (3.25) cannot be used to normalize it. Since governing equation is the Schrödinger equation which is linear and homogeneous in the wave function, the final result is independent of its normalization. For convenience, we can have the following possibilities of normalization:

1. **Flux normalization:**
   In the case of scattering of an incident beam of particles traveling parallel (or nearly parallel), flux of the incident particles is an experimentally measurable quantity (since it is then independent or nearly independent of position) and the wave function can be normalized to unit incident flux in the direction $\hat{k}$

[1] Classical flux of a fluid of density ρ_{cl} and velocity $\vec{v}_{cl}$ is given by the amount of fluid passing through unit area perpendicular to the instantaneous velocity per unit time. This is the fluid contained in a cylinder of unit cross-section and length v_{cl}. Hence flux has magnitude $v_{cl}\rho_{cl}$ in the direction of instantaneous classical velocity.

$$\vec{j}(\vec{r}, t) = 1\,\hat{k}, \tag{3.40}$$

which gives for Eq. (3.39) $A = \frac{1}{\sqrt{v_{in}}}$.

2. **Box normalization:**
Alternately, we can assume the particle to be enclosed in a *large but finite* box of dimension L. For convenience the box is taken to be a cube of sides L, with one corner at the origin and its edges along x-, y- and z-axes. The normalization is done by requiring

$$\int_{\text{cube}} |\psi(\vec{r}, t)|^2 \mathrm{d}^3 r = 1. \tag{3.41}$$

For the wave function (3.39), this gives $A = \frac{1}{\sqrt{L^3}}$.
In this case, since the wave function does not vanish *on the surfaces of the box and outside*, we require that it satisfies *periodic boundary conditions* at corresponding points on opposite walls

$$\begin{aligned} \psi(x, y, z, t) &= \psi(x + L, y, z, t) \\ &= \psi(x, y + L, z, t) \\ &= \psi(x, y, z + L, t). \end{aligned} \tag{3.42}$$

Application of these conditions to the traveling wave, Eq. (3.39) gives

$$\mathrm{e}^{ik_x L} = \mathrm{e}^{ik_y L} = \mathrm{e}^{ik_z L} = 1,$$

Hence

$$k_x = \frac{2\pi n_x}{L}, \quad k_y = \frac{2\pi n_y}{L}, \quad k_z = \frac{2\pi n_z}{L}, \tag{3.43}$$

where n_x, n_y, n_z are integers.
Note that the final results must not depend on the arbitrary length L. Finally we may take the limit $L \to \infty$.
Physical meaning of the box normalization can be understood in the following manner. Imagine the whole space divided in a close pack of cubical boxes with sides (of length L) parallel to the x-, y- and z-axes. With periodic boundary conditions (3.42) the boxes are completely equivalent to each other. Since the overall normalization is arbitrary and the final results do not depend on it, we can normalize the wave function to one within a typical box. Finally taking the limit $L \to \infty$ corresponds to the actual situation.
The wave function (3.39) is an eigen function of momentum with continuous eigen values. Imposition of periodic boundary conditions make the momentum eigen values ($\vec{p} = \hbar\vec{k}$) discrete having Cartesian components

$$p_x = \frac{2\pi\hbar n_x}{L}, \quad p_y = \frac{2\pi\hbar n_y}{L}, \quad p_z = \frac{2\pi\hbar n_z}{L},$$

with n_x, n_y, n_z integers. Discreteness is due to restriction of space within the box. Note that ψ_{TW} (which are momentum eigen functions with eigen value $\hbar\vec{k}$) with different *allowed* values of $\vec{p} = \hbar\vec{k}$ are orthonormal in the box normalization (for brevity we drop the suffix TW, but include the eigen value $\vec{k}$ as a subscript)

$$\int_{\text{cube}} \psi^*_{\vec{k}'}(\vec{r}, t)\psi_{\vec{k}}(\vec{r}, t)\mathrm{d}^3r = \delta_{k_x,k'_x}\delta_{k_y,k'_y}\delta_{k_z,k'_z} = \delta_{n_x,n'_x}\delta_{n_y,n'_y}\delta_{n_z,n'_z}. \tag{3.44}$$

3. **Delta function normalization for continuous eigen values:**
In the mathematical limit $L \to \infty$, the eigen values of components of wave vector ($\vec{k}$), Eq. (3.43), become continuous, since $\delta k_x = \frac{2\pi}{L}\delta n_x$ becomes infinitesimal in the limit $L \to \infty$, with $\delta n_x = 1$ (similarly for other components). Then the Kronecker δ's in Eq. (3.44) should be replaced by the Dirac δ-function (see Sect. 3.4)

$$\int_{\text{cube, } L\to\infty} \psi^*_{\vec{k}'}(\vec{r}, t)\psi_{\vec{k}}(\vec{r}, t)\mathrm{d}^3r = \delta(k_x - k'_x)\delta(k_y - k'_y)\delta(k_z - k'_z) \equiv \delta^{(3)}(\vec{k} - \vec{k}\,'). \tag{3.45}$$

Adjusting the multiplicative constant in $\psi_{\vec{k}}$ to satisfy Eq. (3.45) is called the *delta function normalization* for continuous eigen values. Writing Eq. (3.39) as the momentum eigen function corresponding to momentum $\hbar\vec{k}$

$$\psi_{\vec{k}}(\vec{r}.t) = A\mathrm{e}^{i(\vec{k}.\vec{r}-\omega t)},$$

and using Eqs. (3.45) and (3.38), we have

$$A = \frac{1}{(2\pi)^{\frac{3}{2}}}.$$

Fundamental postulate 2 asserts the completeness of eigen functions [see statements after Eq. (3.1)]. The completeness of the discrete set of momentum eigen functions is given by

$$\sum_{k_x,k_y,k_z} \psi_{\vec{k}}(\vec{r}, t)\,\psi^*_{\vec{k}}(\vec{r}, t) = \sum_{k_x,k_y,k_z} \langle\,\vec{r}\,|\,\psi_{\vec{k}}(t)\,\rangle\langle\,\psi_{\vec{k}}(t)\,|\,\vec{r}\,\rangle = \langle\,\vec{r}\,|\,\vec{r}\,\rangle = 1, \tag{3.46}$$

using closure relation, Eq. (2.49), which in this case is

$$\sum_{k_x,k_y,k_z} |\,\psi_{\vec{k}}(t)\,\rangle\langle\,\psi_{\vec{k}}(t)\,| = 1.$$

The sum Eq. (3.46) is replaced in the $L \to \infty$ limit by an integral over $\vec{k}$, which becomes a continuous variable and replacing $\vec{r}$ by $\vec{r}\,'$ in $\psi^*_{\vec{k}}(\vec{r}, t)$, we have

$$\int \psi_{\vec{k}}(\vec{r}, t)\, \psi^*_{\vec{k}}(\vec{r}\,', t)\, \mathrm{d}^3k = \delta^{(3)}(\vec{r} - \vec{r}\,').$$

This relation follows from Eq. (3.38).

3.6 Problems

1. Consider a time-independent discrete orthonormal basis $\{|\phi_i\rangle, i = 1, 2, \cdots\}$. Obtain the matrix equation in this basis of the time-dependent Schrödinger equation. Also obtain the matrix representations of the coordinate space wave function $|\psi_\alpha(\vec{r}, t)\rangle$ and a time-independent operator $\hat{A}$ in this basis.
2. The state of a particle in a one-dimensional harmonic oscillator well at time $t = 0$ is given by

$$|\psi\rangle = \frac{1}{3}|0\rangle + i\frac{\sqrt{2}}{3}|1\rangle - \frac{\sqrt{6}}{3}|2\rangle,$$

where $|n\rangle$ $(n = 0, 1, 2, \cdots)$ are orthonormal eigen kets of the one-dimensional harmonic oscillator. $|n\rangle$ corresponds to the energy eigen value $(n + \frac{1}{2})\hbar\omega$ (see Chap. 6).

   (a) Is $|\psi\rangle$ normalized?
   (b) What are the possible energy values that can result in an energy measurement? What is the most likely value of energy that will be found in a *single* measurement? What is the probability of finding this value?
   (c) What is the average energy that would result if the energy measurement is repeated many times on identically prepared identical copies of the system? What is the probability of getting this value in a *single* measurement?
   (d) A measurement of energy in a particular case yields the value $\frac{3}{2}\hbar\omega$. The measurement on the system is immediately repeated. What is the result of this measurement? Can one specify the state of the system immediately after this measurement?
   (e) Is the state $|\psi\rangle$ a stationary one? What is the state of the system at time t?

3. Suppose $|\psi_\alpha(x)\rangle$ is an eigen function of a one-dimensional Hamiltonian. Evaluate the quantity $\sum_\alpha \psi^*_\alpha(x)\psi_\alpha(x')$, using bra-ket notation.
4. A particle moves in a complex potential

$$V(\vec{r}, t) = V_0(\vec{r})(1 + i\epsilon),$$

with a small imaginary part $i\epsilon V_0(\vec{r})$ with $V_0(\vec{r})$ real. Obtain the equation of continuity for the probability density and show that the imaginary part corresponds to source or sink of probability, depending on the sign of ϵ. Show that a negative imaginary part causes the total probability to decrease with time, representing absorption.

References

Arfken, G.: Mathematical Methods for Physicists. Academic Press, New York (1966)
Goldstein, H.: Classical Mechanics. Addison-Wesley Publishing Co Inc., Reading, Massachusetts (1950)
Sakurai, J.J. (ed. by Tuan, S.F.): Modern Quantum Mechanics. Addison-Wesley, Delhi (Second Indian reprint (2000)
Schiff, L.I.: Quantum Mechanics, 2nd edn. McGraw-Hill Book Company Inc. (1955). (Reprinted as International Student Edition, Kogakusha Co. Ltd. Tokyo)

Chapter 4
Formulation of Quantum Mechanics: Representations and Pictures

Abstract This chapter introduces different representations for the description of a quantum system: position, momentum and matrix representations and their interrelations. It then goes on to present different pictures for the quantum dynamics: Schrödinger, Heisenberg and interaction pictures from different perspectives. Matrix eigen value equation has been discussed as a mathematical preliminary.

Keywords Position, momentum and matrix representations · Change of representation · Matrix eigen value equation · Schrödinger · Heisenberg and interaction pictures · Heisenberg equation of motion

In this chapter we will discuss formulations of quantum mechanics in different 'representations', the most common of which is the 'position (or coordinate space) representation' for which the fundamental variable is the position. Another useful representation is the 'momentum space representation' for which the fundamental variable is the momentum. The matrix representation leads to Heisenberg's 'matrix mechanics'. Quantum mechanics can also be formulated in terms of an integral equation, called 'Lippmann–Schwinger equation'. The Schrödinger equation in all these representations are equivalent and are different projections of the Hilbert space Eq. (3.21).

We can visualize the motion of the system, i.e. its time evolution, in two different perspectives: we can let the 'state' evolve in time, while the 'operators' remain stationary, resulting in the so-called Schrödinger picture. Alternatively, we can allow the state to remain stationary, while the operators evolve in time. This approach gives rise to the 'Heisenberg picture'. Both formulations will lead to the same result for observables, which are basically the expectation values of operators. Since the connection with the physical world is made through observables, it is natural that these two approaches will give the same final result.

4.1 Position (Coordinate) Representation

We saw in Chap. 3 that the position operator ($\hat{\vec{r}}$) is a multiplying function in position (coordinate) representation. Then in order that the position-momentum uncertainty relation is satisfied, the momentum operator has to be represented by the differential

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5_4

operator $\hat{\vec{p}} \to -i\hbar\vec{\nabla}$. In Chap. 3, Sect. 3.3, we discussed the eigen value equation of the position operator in coordinate space. Here we briefly recapitulate the basic equations. According to the postulates of quantum mechanics, the position operator ($\hat{\vec{r}}$), being an observable, must be Hermitian. It satisfies an eigen value equation

$$\hat{\vec{r}}\,|\vec{r}\,'\rangle = \vec{r}\,'|\vec{r}\,'\rangle.$$

The set of eigen kets $\{|\vec{r}\,'\rangle\}$ forms a complete set. The eigen kets are assumed to be normalized. Because the eigen values are continuous, the Kronecker delta and sum in the orthonormality and completeness relations of Chap. 2 are to be replaced by Dirac delta function and integral respectively

$$\langle\phi_i|\phi_j\rangle = \delta_{ij} \quad \text{replaced by} \to \quad \langle\vec{r}\,'|\vec{r}\,''\rangle = \delta^{(3)}(\vec{r}\,' - \vec{r}\,''). \tag{4.1}$$

$$\sum_i |\phi_i\rangle\langle\phi_i| = \hat{1} \quad \text{replaced by} \to \quad \int |\vec{r}\,'\rangle\, \mathrm{d}^3r\,'\langle\vec{r}\,'| = \hat{1}, \tag{4.2}$$

We also saw that the state $|\psi(t)\rangle$ can be expanded in the CONS of position eigen kets $\{|\vec{r}\,'\rangle\}$

$$\begin{aligned} |\psi(t)\rangle &= \int |\vec{r}\,'\rangle\, \mathrm{d}^3r\,'\langle\vec{r}\,'|\psi(t)\rangle \\ &\equiv \int \psi(\vec{r}\,', t)|\vec{r}\,'\rangle\, \mathrm{d}^3r\,', \end{aligned} \tag{4.3}$$

where $\psi(\vec{r}\,', t) = \langle\vec{r}\,'|\psi(t)\rangle$. This equation gives rise to the interpretation that $|\psi(\vec{r}, t)|^2$ is the position probability density (see Sect. 3.3, Chap. 3) and is normalized according to

$$\int |\psi(\vec{r}, t)|^2 \mathrm{d}^3r = 1 \qquad \text{[Eq. (3.25)]}.$$

Schrödinger equation in coordinate space

Let us now see how the Schrödinger equation in the abstract Hilbert space looks in the coordinate space. Taking inner product of the time-dependent Schrödinger equation, (3.21) with $\langle\vec{r}\,|$, we have

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\langle\vec{r}\,|\psi(t) &= \langle\vec{r}\,|\hat{H}|\psi(t)\rangle \\ \text{or,} \quad i\hbar\frac{\partial}{\partial t}\psi(\vec{r}, t) &= \int \langle\vec{r}\,|\hat{H}|\vec{r}\,'\rangle\, \mathrm{d}^3r\,'\langle\vec{r}\,'|\psi(t)\rangle \\ &= \int H(\vec{r}, t)\ \delta^{(3)}(\vec{r} - \vec{r}\,')\, \mathrm{d}^3r\,'\ \langle\vec{r}\,'|\psi(t)\rangle \\ &= H(\vec{r}, t)\psi(\vec{r}, t) \\ &= \Big[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r}, t)\Big]\psi(\vec{r}, t). \end{aligned} \tag{4.4}$$

Here we have taken the Hamiltonian to be a 'local' operator, i.e.

$$\langle \vec{r}\,|\hat{H}(t)|\vec{r}\,'\rangle = H(\vec{r},t)\,\delta^{(3)}(\vec{r}-\vec{r}\,').$$

Equation (4.4) is the coordinate space time-dependent Schrödinger equation obtained by projection of the Hilbert space Eq. (3.21) on to the coordinate space.

Next consider two vectors $|\alpha\rangle$ and $|\beta\rangle$ in the Hilbert space $\mathcal{H}$, representing two states of the system at time t. Using Eq. (4.2), the inner product is (suppressing time dependence for brevity)

$$\langle\beta|\alpha\rangle = \langle\beta|\int|\vec{r}\,\rangle \mathrm{d}^3r\,\langle\vec{r}\,|\alpha\rangle = \int\langle\beta|\vec{r}\,\rangle\langle\vec{r}\,|\alpha\rangle \mathrm{d}^3r = \int\psi^*_\beta(\vec{r}\,)\psi_\alpha(\vec{r}\,)\mathrm{d}^3r, \tag{4.5}$$

where $\psi_\alpha(\vec{r}\,)$ and $\psi_\beta(\vec{r}\,)$ are the coordinate space wave functions of the states $|\alpha\rangle$ and $|\beta\rangle$ respectively, according to definition (4.3). Extreme right side of Eq. (4.5) is the usual definition of overlap integral, and if $\psi_\beta(\vec{r}\,)$ and $\psi_\alpha(\vec{r}\,)$ are taken as vectors in $\mathcal{L}^2$ (vector space of square integrable functions), then this is the definition of inner product. Thus the inner product $\langle\beta|\alpha\rangle$ is independent of the representation. It is also the probability amplitude for the state $|\beta\rangle$ to be found in the state $|\alpha\rangle$.

We can now obtain the coordinate space representation of an arbitrary matrix element of the Hilbert space operator $\hat{A}$. The matrix formed by the numerical values of the matrix elements is the matrix representation of the operator $\hat{A}$. Introducing Eq. (4.2) twice, we have

$$\begin{aligned} A_{\alpha\beta} \equiv \langle\alpha|\hat{A}|\beta\rangle &= \int\int\langle\alpha|\vec{r}\,'\rangle\,\mathrm{d}^3r\,'\langle\vec{r}\,'|\hat{A}|\vec{r}\,\rangle\,\mathrm{d}^3r\,\langle\vec{r}\,|\beta\,\rangle \\ &= \int\int\psi^*_\alpha(\vec{r}\,')A(\vec{r}\,',\vec{r}\,)\psi_\beta(\vec{r}\,)\,\mathrm{d}^3r\,'\mathrm{d}^3r \\ &= \int\psi^*_\alpha(\vec{r}\,)A(\vec{r}\,)\psi_\beta(\vec{r}\,)\,\mathrm{d}^3r, \end{aligned} \tag{4.6}$$

where we assume the operator $\hat{A}$ to be local, $A(\vec{r}\,',\vec{r}\,) = A(\vec{r}\,)\,\delta^{(3)}(\vec{r}\,'-\vec{r}\,)$.

Next we discuss how the eigen value equation of a Hermitian operator looks in the coordinate representation. Consider an observable $\hat{A}$ on $\mathcal{H}$, satisfying an eigen value equation, assuming that the eigen vectors are orthonormalized

$$\hat{A}|\phi_i\rangle = a_i|\phi_i\rangle. \tag{4.7}$$

We can expand an arbitrary state $|\alpha\rangle$ in the complete orthonormal set $\{|\phi_i\rangle\}$

$$|\alpha\rangle = \sum_i c_i|\phi_i\rangle.$$

Taking inner product with $\langle \vec{r}\,|$ on both sides

$$\langle \vec{r}\,|\alpha\rangle = \sum_i c_i \langle \vec{r}\,|\phi_i\rangle.$$

Thus we have, using the definition (4.3),

$$\psi_\alpha(\vec{r}\,) = \sum_i c_i \phi_i(\vec{r}\,). \tag{4.8}$$

Next we obtain the projection of the eigen value Eq. (4.7) by taking its inner product with $\langle \vec{r}\,|$ and introducing Eq. (4.2)

$$\langle \vec{r}\,|\hat{A}\int |\vec{r}\,'\rangle\, \mathrm{d}^3 r'\, \langle \vec{r}\,'|\phi_i\rangle = a_i\, \langle \vec{r}\,|\phi_i\rangle.$$

Assuming $\hat{A}$ to be local, i.e. $\langle \vec{r}\,|\hat{A}|\vec{r}\,'\rangle = A(\vec{r}\,)\,\delta^{(3)}(\vec{r}-\vec{r}\,')$, we have

$$A(\vec{r}\,)\,\phi_i(\vec{r}\,) = a_i\,\phi_i(\vec{r}\,). \tag{4.9}$$

We now have the eigen value equation of the operator $\hat{A}(\vec{r}\,)$ in coordinate representation, where $\phi_i(\vec{r}\,)$ is the eigen function corresponding to the eigen value a_i. Note that the set of eigen values remains the same. The operator $\hat{A}(\vec{r}\,)$ may involve differential operators (as in coordinate representation of momentum operator $\hat{\vec{p}} = -i\hbar\vec{\nabla}$). In that case, Eq. (4.9) is a differential eigen value equation in $\vec{r}$. Equation (4.8) is equivalent to the expansion of the 'wave function' $\psi_\alpha(\vec{r}\,)$ in the complete set of eigen functions $\{\phi_i(\vec{r}\,)\}$ of the coordinate space operator $A(\vec{r}\,)$.

We can verify that the set of all eigen functions of the Schrödinger equation forms a vector space of square integrable functions of the appropriate coordinates (this vector space is called Hilbert space, see Sect. 2.9, Chap. 2). This is consistent with Postulate 1 of quantum mechanics (see Sect. 3.2, Chap. 3).

4.2 Momentum Representation

In momentum space representation, the momentum operator is the multiplying function, while the position operator becomes a differential operator on the variable $\vec{p}$. The uncertainty relation is satisfied when position operator is represented by $\hat{\vec{r}} \to i\hbar\vec{\nabla}_{\vec{p}}$, where the gradient operator $\vec{\nabla}_{\vec{p}}$ acts on the vector variable $\vec{p}$. As in the case of coordinate representation, for momentum representation we first look for the eigen state $|\,\vec{p}\,'\rangle$ of the momentum operator corresponding to momentum eigen value $\vec{p}\,'$

$$\hat{\vec{p}}\,|\,\vec{p}\,'\rangle = \vec{p}\,'\,|\,\vec{p}\,'\rangle \tag{4.10}$$

The set of all eigen vectors $\{|\,\vec{p}\,\rangle\}$ is complete and assumed to be orthonormal with continuous eigen value $\vec{p}$. Since momentum is a continuous variable, the orthonormality property is given by

$$\langle\,\vec{p}\,'|\,\vec{p}\,''\rangle = \delta^{(3)}(\vec{p}\,' - \vec{p}\,''), \tag{4.11}$$

and the closure (completeness) is given by

$$\int |\,\vec{p}\,'\rangle\,\mathrm{d}^3p\,'\,\langle\,\vec{p}\,'| = \hat{1}. \tag{4.12}$$

An arbitrary state $|\alpha\rangle$ (suppressing its time dependence) can be expanded in the complete orthonormal set $\{|\,\vec{p}\,\rangle\}$, using Eq. (4.12)

$$|\alpha\rangle = \int |\,\vec{p}\,\rangle\,\mathrm{d}^3p\,\langle\,\vec{p}\,|\alpha\rangle \equiv \int \phi_\alpha(\,\vec{p}\,)\,|\,\vec{p}\,\rangle\,\mathrm{d}^3p.$$

The expansion coefficient $\phi_\alpha(\,\vec{p}\,)$ (which is a function of $\vec{p}\,$) is called the 'momentum space wave function' of the state $|\alpha\rangle$. According to Corollary 2b of Chap. 3, Sect. 3.2, $|\phi_\alpha(\,\vec{p}\,)|^2\,\mathrm{d}^3p$ is the probability of finding the particle with momentum lying within a small volume element d^3p about the value $\vec{p}\,$. If the state $|\alpha\rangle$ is normalized, i.e. $\langle\alpha|\alpha\rangle = 1$, then introducing the closure relation (4.12)

$$\int \langle\alpha|\,\vec{p}\,\rangle\,\mathrm{d}^3p\,\langle\,\vec{p}\,|\alpha\rangle = 1,$$

or,

$$\int |\phi_\alpha(\,\vec{p}\,)|^2\,\mathrm{d}^3p\, = 1,$$

giving the normalization of the momentum space wave function. This is consistent with the probability interpretation given above.

In general, we will denote momentum space wave functions by $\phi(\vec{p})$, while coordinate space wave function by $\psi(\vec{r}\,)$. However the real identification of the type of wave function is by the argument (position or momentum variable) of the wave function.

Schrödinger equation in momentum space

We can obtain the Schrödinger equation in momentum space from the abstract Schrödinger equation in the Hilbert space, in a manner similar to the coordinate space case. The result is

$$i\hbar\frac{\partial}{\partial t}\phi(\vec{p},t) = \Big[\frac{p^2}{2m} + V(i\hbar\vec{\nabla}_{\vec{p}}\,,t)\Big]\phi(\vec{p},t), \tag{4.13}$$

provided $V(\vec{r}, t)$ is an analytic function of $\vec{r}$, such that $V(i\hbar\vec{\nabla}_{\vec{p}}, t)$ is the potential function expressed in terms of momentum operator (replacing $\vec{r}$ by $i\hbar\vec{\nabla}_{\vec{p}}$).

Momentum eigen function in coordinate space

We can calculate the coordinate space eigen function of the momentum operator by taking the inner product of the momentum eigen value Eq. (4.10) with a position bra $\langle \vec{r} \,|$

$$\langle \vec{r} \,|\hat{\vec{p}}\,|\, \vec{p}\,' \rangle = \vec{p}\,' \langle \vec{r} \,|\, \vec{p}\,' \rangle.$$

Next, introducing closure relation in coordinate space, Eq. (4.2)

$$\int \langle \vec{r} \,|\hat{\vec{p}}\,|\, \vec{r}\,' \rangle \, \mathrm{d}^3 r' \, \langle \vec{r}\,' \,|\, \vec{p}\,' \rangle = \vec{p}\,' \langle \vec{r} \,|\, \vec{p}\,' \rangle.$$

Now $\langle \vec{r} \,|\hat{\vec{p}}\,|\, \vec{r}\,' \rangle = \hat{\vec{p}}\, \delta^{(3)}(\vec{r} - \vec{r}\,')$, and writing $\psi_{\vec{p}\,'}(\vec{r}) = \langle \vec{r} \,|\, \vec{p}\,' \rangle$, we have

$$\hat{\vec{p}}\, \psi_{\vec{p}\,'}(\vec{r}) = \vec{p}\,' \psi_{\vec{p}\,'}(\vec{r}).$$

Since in coordinate space, $\hat{\vec{p}} = -i\hbar\vec{\nabla}$, we have

$$\vec{\nabla} \psi_{\vec{p}\,'}(\vec{r}) = \frac{i}{\hbar} \vec{p}\,' \psi_{\vec{p}\,'}(\vec{r}).$$

Its solution is

$$\psi_{\vec{p}\,'}(\vec{r}) = C \exp\left(\frac{i}{\hbar} \vec{p}\,' . \vec{r}\right).$$

The normalization constant C is obtained from Eq. (4.11)

$$\begin{aligned}
\delta^{(3)}(\vec{p}\,' - \vec{p}\,'') = \langle \vec{p}\,' \,|\, \vec{p}\,'' \rangle &= \int \langle \vec{p}\,' \,|\, \vec{r} \rangle \, \mathrm{d}^3 r \, \langle \vec{r} \,|\, \vec{p}\,'' \rangle = \int \psi^*_{\vec{p}\,'}(\vec{r}) \, \psi_{\vec{p}\,''}(\vec{r}) \, \mathrm{d}^3 r \\
&= |C|^2 \int \exp\Big(\frac{i}{\hbar} (\vec{p}\,'' - \vec{p}\,') \cdot \vec{r}\Big) \mathrm{d}^3 r = |C|^2 \hbar^3 \int \exp\Big(i \, (\vec{p}\,'' - \vec{p}\,') \cdot \vec{\rho}\Big) \mathrm{d}^3 \rho \\
&= |C|^2 \, \hbar^3 (2\pi)^3 \, \delta^{(3)} \, (\vec{p}\,' - \vec{p}\,''),
\end{aligned}$$

using Eq. (3.38) and noting that the delta function is symmetric in its argument. Hence $|C| = (2\pi\hbar)^{-3/2}$, and we have

$$\psi_{\vec{p}}\,(\vec{r}) = (2\pi\hbar)^{-3/2} \exp\left(\frac{i}{\hbar}\, \vec{p} \cdot \vec{r}\right). \tag{4.14}$$

This is the coordinate space eigen function of the momentum operator for a fixed value of $\vec{p}$. It represents a plane wave with constant amplitude and momentum eigen value $\vec{p}$. Note that $|\psi_{\vec{p}}(\vec{r})|^2$ is a constant. Hence the position uncertainty is

infinite for a particle with a definite momentum $\vec{p}$, in agreement with the uncertainty principle.

Similarly, the eigen function of position corresponding to eigen value $\vec{r}$ in momentum space is

$$\phi_{\vec{r}}(\vec{p}) = (2\pi\hbar)^{-3/2} \exp\Big(-\frac{i}{\hbar}\,\vec{p}\cdot\vec{r}\Big). \tag{4.15}$$

4.3 Change of Representation

We can express the wave function in one representation into the wave function in another representation, using a suitable closure relation. As an example, suppose the momentum space wave function $\phi_\alpha(\vec{p})$ is known and it is desired to express it in the coordinate representation. For this purpose, insert the closure relation for momentum eigen states, Eq. (4.12) and use momentum eigen function in coordinate space, Eq. (4.14)

$$\begin{aligned}\psi_\alpha(\vec{r}) &= \langle\vec{r}\,|\alpha\rangle = \int\langle\vec{r}\,|\,\vec{p}\,\rangle\,\mathrm{d}^3p\,\langle\vec{p}\,|\alpha\rangle = \int\psi_{\vec{p}}(\vec{r})\phi_\alpha(\vec{p})\,\mathrm{d}^3p \\ &= (2\pi\hbar)^{-3/2}\int\phi_\alpha(\vec{p})\exp\Big(\frac{i}{\hbar}\,\vec{p}\cdot\vec{r}\Big)\,\mathrm{d}^3p.\end{aligned} \tag{4.16}$$

In mathematical language, this is called the Fourier integral representation of the coordinate space wave function $\psi_\alpha(\vec{r})$. As we see below, $\psi_\alpha(\vec{r})$ is also called the inverse Fourier transform (IFT) of the momentum space wave function $\phi_\alpha(\vec{p})$ (Chattopadhyay 2006).

In a similar fashion, using the closure relation of coordinate eigen states and Eq. (4.15) we get

$$\phi_\alpha(\vec{p}) = \langle\vec{p}\,|\alpha\rangle = (2\pi\hbar)^{-3/2}\int\psi_\alpha(\vec{r})\exp\Big(-\frac{i}{\hbar}\,\vec{p}\cdot\vec{r}\Big)\,\mathrm{d}^3r, \tag{4.17}$$

which is called the Fourier transform (FT) of $\psi_\alpha(\vec{r})$.

Mathematical preliminary: Fourier transform and inverse transform

The set of functions $\{(2\pi)^{-3/2}\,\mathrm{e}^{i\vec{k}\cdot\vec{r}}\}$ which are solutions of Hermitian eigen value equation $-\nabla^2\phi_{\vec{k}}(\vec{r}) = k^2\phi_{\vec{k}}(\vec{r})$, as a function of the continuous three-dimensional variable $\vec{k}$, forms a complete orthonormal set of functions over all space, see Eq. (3.38). Hence any function, $f(\vec{r})$, which is analytic (or at least piecewise analytic) in the entire interval, can be expanded in this set (as an integral, since the eigen value is continuous)

$$f(\vec{r}) = (2\pi)^{-3/2}\int g(\vec{k})\,\mathrm{e}^{i\vec{k}\cdot\vec{r}}\,\mathrm{d}^3k.$$

This representation of $f(\vec{r})$ is called Fourier integral representation, and the function $g(\vec{k})$ is called the Fourier transform (FT) of $f(\vec{r})$. As we have seen above, $g(\vec{k})$ can be expressed as

$$g(\vec{k}) = (2\pi)^{-3/2} \int f(\vec{r})\, \mathrm{e}^{-i\vec{k}\cdot\vec{r}}\, \mathrm{d}^3k,$$

The function $f(\vec{r})$ is called the inverse Fourier transform (IFT) of $g(\vec{k})$. Here we follow the definitions of FT and IFT according to Ref. Chattopadhyay (2006). Note that substitution of one into the other should give back the original function, using Eqs. (3.38) and (3.37). This requires that the argument of exponential functions in FT and IFT should have *opposite* signs (see Chap. 3, Sect. 3.4). Thus the definition of FT is flexible up to the sign of argument of exponential function (once it is chosen, the expression for IFT is fixed). We choose negative sign of the argument of the exponential function in the definition of FT. The opposite is used in Ref. Arfken (1966).

In one dimension the Fourier integral representation of $f(x)$ is

$$f(x) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} g(k)\, \mathrm{e}^{ikx}\, \mathrm{d}k,$$

while $g(k)$ is given by

$$g(k) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} f(x)\, \mathrm{e}^{-ikx}\, \mathrm{d}x.$$

$g(k)$ is called the Fourier transform, while $f(x)$ is the inverse transform.

4.4 Matrix Representation: Matrix Mechanics

According to the *fundamental postulate 1*, a physical system is associated with a Hilbert space $\mathcal{H}$ and any vector $|\psi(t)\rangle$ belonging to it is a possible realizable state of the system at time t. From our discussion of vector spaces, we saw that a *discrete*, time-independent orthonormal basis, $\{|\phi_i\rangle, i = 1, 2, \cdots\}$, can always be chosen to span $\mathcal{H}$, and is thus *complete*. The basis is finite or infinite depending on the vector space being finite or infinite dimensional. Then the state vector $|\psi(t)\rangle$ can be expanded in this complete orthonormal set (CONS)

$$|\psi(t)\rangle = \sum_i a_i(t)|\phi_i\rangle,$$

where $a_i(t) = \langle \phi_i | \psi(t) \rangle$. Then the column matrix

$$\underline{a}(t) = \begin{pmatrix} a_1(t) \\ a_2(t) \\ \cdots \end{pmatrix}$$

represents the state $|\psi(t)\rangle$. The Schrödinger equation is

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H}(t) |\psi(t)\rangle.$$

Taking inner product with $\langle \phi_i |$ and introducing the closure relation, Eq. (2.49) we get

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} a_i(t) &= \sum_j \langle \phi_i | \hat{H}(t) | \phi_j \rangle \langle \phi_j | \psi(t) \rangle \\ &= \sum_j H_{ij}(t) a_j(t). \end{aligned}$$

This can be written as a matrix equation

$$i\hbar \frac{\partial}{\partial t} \underline{a}(t) = \underline{\underline{H}}(t)\, \underline{a}(t), \tag{4.18}$$

where $\underline{\underline{H}}(t)$ is the square Hamiltonian matrix in our chosen time-independent basis. Thus we get the time-dependent Schrödinger equation in matrix form, satisfied by the Hamiltonian matrix in $\{|\phi_i\rangle\}$ basis. The column vector $\underline{a}(t)$ determines the state vector $|\psi(t)\rangle$. Equation (4.18) is equivalent to the Schrödinger equation but now deals with a set of numbers, which in general depend on time. If $\hat{H}$ is independent of t, time dependence can be factored out as

$$\underline{a}(t) = \underline{a}\, T(t). \tag{4.19}$$

Substituting in Eq. (4.18), we have

$$\frac{1}{T(t)} \left(i\hbar \frac{\mathrm{d}T(t)}{\mathrm{d}t} \right) \underline{a} = \underline{\underline{H}}\, \underline{a}.$$

Right side is independent of t, hence the factor in front of column vector $\underline{a}$ on the left side must be a constant (independent of t), say, E. Then

$$\underline{\underline{H}}\, \underline{a} = E\, \underline{a}, \tag{4.20}$$

and

$$i\hbar \frac{\mathrm{d}T(t)}{\mathrm{d}t} = E\, T(t),$$

Solution of the last equation is

$$T(t) = \mathrm{e}^{-iEt/\hbar}, \tag{4.21}$$

leaving out a constant multiplying factor, which can be absorbed in $\underline{a}$. From Eq. (4.20), we see that $\underline{\underline{H}}$ satisfies a matrix eigen value equation. There will be a number of real (since $\underline{\underline{H}}$ is Hermitian) eigen values, $E_1, E_2, \cdots$ corresponding to eigen column vectors $\underline{a_1}, \underline{a_2}, \cdots$ respectively:

$$\underline{\underline{H}}\,\underline{a_n} = E_n\,\underline{a_n}, \quad (n = 1, 2, \cdots). \tag{4.22}$$

Next, we consider the expectation value of an observable ω, represented by a time-independent operator $\hat{\mathcal{O}}$ acting on the Hilbert space $\mathcal{H}$. Let $\hat{\mathcal{O}}$ satisfy a discrete eigen value equation

$$\hat{\mathcal{O}}\,|\zeta_i\rangle = \omega_i\,|\zeta_i\rangle.$$

In our chosen basis $\{|\phi_i\rangle, i = 1, 2, \cdots\}$ this equation becomes

$$\sum_j \langle\phi_k|\hat{\mathcal{O}}|\phi_j\rangle\langle\phi_j|\zeta_i\rangle = \omega_i\langle\phi_k|\zeta_i\rangle,$$

where we projected on $\langle\phi_k|$ and inserted the closure relation. In matrix notation, this can be written as

$$\sum_j O_{kj}\,b_{ji} = \omega_i\,b_{ki},$$

where $O_{kj} = \langle\phi_k|\hat{\mathcal{O}}|\phi_j\rangle$ and $b_{ki} = \langle\phi_k|\zeta_i\rangle$. The number O_{kj} is the (kj)th element of the matrix $\underline{\underline{O}}$ in the chosen basis. For a fixed i, the numbers b_{ki}, $(k = 1, 2, \cdots)$ constitute the elements of a column vector $\underline{b_i}$, which is the eigen vector of the Hermitian matrix $\underline{\underline{O}}$, corresponding to real eigen value ω_i. Now the expectation value of the observable ω in the state represented by $|\psi\rangle$ is, according to Eq. (3.3)

$$\begin{aligned}\langle\hat{\mathcal{O}}\rangle &= \frac{\langle\psi|\hat{\mathcal{O}}|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\sum_{ij}\langle\psi|\phi_i\rangle\langle\phi_i|\hat{\mathcal{O}}|\phi_j\rangle\langle\phi_j|\psi\rangle}{\sum_i\langle\psi|\phi_i\rangle\langle\phi_i|\psi\rangle} \\ &= \frac{\sum_{ij} a_i^*\,O_{ij}\,a_j}{\sum_i a_i^*\,a_i} = \frac{\underline{a}^\dagger\,\underline{\underline{O}}\,\underline{a}}{\underline{a}^\dagger\underline{a}},\end{aligned} \tag{4.23}$$

where $\underline{a}^\dagger$ is the row matrix formed by writing complex conjugate of the elements of the column matrix $\underline{a}$, in a row.

Equation (4.22) is a matrix eigen value equation, involving only numbers arranged in matrix form. It is equivalent to the time-independent Schrödinger equation and the

elements of the eigen column vector, together with its time dependence, Eqs. (4.19) and (4.21) define the state vector $|\psi(t)\rangle$. If the Hamiltonian is time-dependent, then Eq. (4.18) is to be solved. This equation is again a matrix equation involving numbers only, although the numbers are now time-dependent and the equation involves a time derivative. The expectation value of an observable is obtained in terms of the resultant numbers in $\underline{a}$. A square and a column matrix correspond respectively to an operator (for an observable the matrix is Hermitian) and a state ket. The row matrix $\underline{a}^\dagger$ corresponds to a bra vector. Hence by rules of matrix multiplication, $\underline{a}^\dagger \underline{\underline{O}}\, \underline{a}$ as well as $\underline{a}^\dagger \underline{a}$ are ordinary numbers (which may be time-dependent).

The column matrices form a (ket) vector space. Alternatively, the set of (in general complex) numbers (which are the elements of the column matrix) can be considered as a vector space of multi-tuple numbers. We saw in Sect. 2.9, Chap. 2 that such entities form a Hilbert space, in agreement with Postulate 1 of Sect. 3.2, Chap. 3.

Thus a physical problem can be solved completely using matrices only. This procedure, originally proposed by Heisenberg, is called the *matrix mechanics*. As we see from the foregoing discussions, it is completely equivalent to Schrödinger's wave function approach, as also the abstract vector space approach starting from the Schrödinger Eq. (3.21) in abstract Hilbert space. An example of the last approach will be discussed in Chap. 6.

4.5 Math-Prelim: Matrix Eigen Value Equation

In this section, we discuss how the matrix eigen value Eq. (4.22) can be solved to get the eigen values and corresponding eigen vectors. This equation can be written as

$$(\underline{\underline{H}} - E_n \underline{\underline{1}})\, \underline{a}_n = \underline{0}, \tag{4.24}$$

where $\underline{0}$ is the null column matrix and $\underline{\underline{1}}$ is the unit matrix, whose diagonal elements are 1 and all off-diagonal elements vanish

$$\underline{\underline{1}}_{ij} = \delta_{ij}.$$

The column vector corresponding to the eigen value E_n, i.e. $\underline{a}_n$, has components $a_{1n}, a_{2n}, \cdots$. Hence Eq. (4.24) reads in long hand

$$\begin{aligned}
&(\underline{\underline{H}} - E_n \underline{\underline{1}})(\underline{a})_n \\
&= \begin{pmatrix}
(H_{11} - E_n) & H_{12} & \cdots & H_{1j} & \cdots \\
H_{21} & (H_{22} - E_n) & \cdots & H_{2j} & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
H_{j1} & H_{j2} & \cdots & (H_{jj} - E_n) & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots
\end{pmatrix}
\begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{jn} \\ \cdots \end{pmatrix}
= \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 0 \\ \cdots \end{pmatrix}.
\end{aligned}$$

This is a set of linear homogeneous equations (LHE) for the unknowns $\{a_{1n}, a_{2n}, \cdots\}$. For a non-trivial solution, the determinant of the coefficients must vanish

$$\begin{vmatrix} (H_{11}-E_n) & H_{12} & \cdots & H_{1j} & \cdots \\ H_{21} & (H_{22}-E_n) & \cdots & H_{2j} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ H_{j1} & H_{j2} & \cdots & (H_{jj}-E_n) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \end{vmatrix} = 0.$$

This is called the *secular equation* for the unknown E_n. In general the Hilbert space is infinite dimensional. Consequently the set of LHE and the determinant are infinite dimensional. For a practical calculation these must be truncated. The basis having successive members $|\phi_i\rangle$ with $i = 1, 2, \cdots$ is so chosen that the expansion of $|\psi(t)\rangle$ in this basis converges fast. Suppose the desired convergence is reached at a value N_M of i. This means that the basis set can be truncated to $\{|\phi_1\rangle, |\phi_2\rangle, \cdots, |\phi_{N_M}\rangle\}$. Consequently, the matrix Eq. (4.24) is of dimension N_M. Then we have a set of N_M equations and an $N_M \times N_M$ determinant

$$\begin{vmatrix} (H_{11}-E_n) & H_{12} & \cdots & H_{1N_M} \\ H_{21} & (H_{22}-E_n) & \cdots & H_{2N_M} \\ \cdots & \cdots & \cdots & \cdots \\ H_{N_M 1} & H_{N_M 2} & \cdots & (H_{N_M N_M}-E_n) \end{vmatrix} = 0.$$

With known elements of the Hamiltonian matrix, this is a polynomial (of degree N_M) equation for the unknown E_n. Hence there are N_M solutions for E_n, say, $E_1, E_2, \cdots, E_{N_M}$. These are the eigen values. For a particular eigen value E_k, the matrix equation for the corresponding column matrix $(\underline{a})_k$ becomes

$$(\underline{\underline{H}} - E_k \underline{\underline{1}})(\underline{a})_k = \begin{pmatrix} (H_{11}-E_k) & H_{12} & \cdots & H_{1N_M} \\ H_{21} & (H_{22}-E_k) & \cdots & H_{2N_M} \\ \cdots & \cdots & \cdots & \cdots \\ H_{N_M 1} & H_{N_M 2} & \cdots & (H_{N_M N_M}-E_k) \end{pmatrix} \begin{pmatrix} a_{1k} \\ a_{2k} \\ \cdots \\ a_{N_M k} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 0 \end{pmatrix}.$$

For a particular selected value of E_k, this is a set of linear homogeneous equations for the unknowns $a_{1k}, a_{2,k}, \cdots, a_{N_M k}$. A non-trivial solution is guaranteed, since the selected value of E_k makes the determinant of the coefficients vanish. However the set of equations being homogeneous, these unknowns are arbitrary up to an overall normalization constant. To solve, we take one of them equal to C (say $a_{1k} = C$) in the above equation and solve for the rest, $a_{2k}, \cdots, a_{N_M k}$, in terms of C. Finally calculate C by normalization

$$\sum_{j=1}^{j=N_M} |a_{jk}|^2 = 1.$$

Thus we get the normalized eigen vector $(\underline{a})_k$ corresponding to eigen value E_k. This process is to be repeated for $k = 1, \cdots, N_M$.

In practice, the process of truncation has to be repeated with increasing N_M, until the desired eigen values (usually the lowest ones) become convergent, up to a pre-set desired limit. For a quicker convergence (i.e. a smaller N_M), one has to apply physics intuition and previous knowledge of same or similar problems to choose the truncated basis $\{|\phi_1\rangle, |\phi_2\rangle, \cdots, |\phi_{N_M}\rangle\}$ (including the ordering of the basis elements) to obtain the Hamiltonian matrix $\underline{\underline{H}}$.

Note that if E_n is not one of the eigen values, the set of LHE has no non-trivial solutions, so the corresponding column vector vanishes. This is a trivial solution of Eq. (4.24).

4.6 Quantum Dynamics—Perspectives: Schrödinger, Heisenberg and Interaction Pictures

Note that abstract mathematical entities like state vectors and operators in the Hilbert space are not directly accessible to measurement. The measurable quantities are the eigen values, expansion coefficients (inner products) and expectation values. Measurement of an observable $\hat{A}$ corresponds to getting one of its eigen values a_i. The probability of getting a particular eigen value a_i (corresponding orthonormalized eigen ket being $|a_i\rangle$) for the system in the state $|\psi\rangle$ (assumed normalized, $\langle\psi|\psi\rangle = 1$), is $|\langle a_i|\psi\rangle|^2$. The average result of measurements of $\hat{A}$ in the state $|\psi\rangle$ is the expectation value $\langle\psi|\hat{A}|\psi\rangle$, (which is also an inner product of $\langle\psi|$ and $\hat{A}|\psi\rangle$). Thus any formulation of quantum mechanics is equally acceptable, if measurable qualities remain unchanged in that formulation.

We can see that under a unitary transformation these quantities remain unchanged. Consider a unitary transformation $\hat{\mathcal{U}}$, which transforms a ket $|\alpha\rangle$ to $|\alpha\rangle'$ (suppressing dependence on time, assuming that all quantities are at a particular time)

$$|\alpha\rangle' = \hat{\mathcal{U}}|\alpha\rangle \quad (\text{with } \hat{\mathcal{U}}^\dagger\hat{\mathcal{U}} = \hat{\mathcal{U}}\hat{\mathcal{U}}^\dagger = \hat{1}) \tag{4.25}$$

Then under this transformation, inner products remain unchanged

$$\begin{aligned} {}'\langle\beta|\alpha\rangle' &= \langle\beta|\hat{\mathcal{U}}^\dagger\hat{\mathcal{U}}|\alpha\rangle \\ &= \langle\beta|\alpha\rangle \end{aligned}$$

Suppose the Hermitian operator $\hat{A}$ satisfies an eigen value equation

$$\hat{A}|a_i\rangle = a_i|a_i\rangle \quad (i = 1, 2, \cdots.)$$

Premultiply by $\hat{\mathcal{U}}$ and introduce $\hat{\mathcal{U}}^\dagger\hat{\mathcal{U}} = \hat{1}$

$$\hat{\mathcal{U}}\hat{A}(\hat{\mathcal{U}}^\dagger\hat{\mathcal{U}})|a_i\rangle = a_i\hat{\mathcal{U}}|a_i\rangle$$
$$\text{or,} \qquad (\hat{\mathcal{U}}\hat{A}\hat{\mathcal{U}}^\dagger)|a_i\rangle' = a_i|a_i\rangle'.$$

Thus $|a_i\rangle'$ will be the eigen vector corresponding to the *same* eigen value a_i for the operator $\hat{\mathcal{U}}\hat{A}\hat{\mathcal{U}}^\dagger$ (which can easily be shown to be Hermitian). Thus under this unitary transformation an operator transforms as

$$\hat{A} \to \hat{A}' = \hat{\mathcal{U}}\hat{A}\hat{\mathcal{U}}^\dagger, \tag{4.26}$$

such that

$$\hat{A}'|a_i\rangle' = a_i|a_i\rangle'.$$

We can also verify that the expectation value of $\hat{A}$ in the state $|\psi\rangle$ remains unchanged under the unitary transformation

$$'\langle\psi|\hat{A}'|\psi\rangle' = \langle\psi|\hat{\mathcal{U}}^\dagger(\hat{\mathcal{U}}\hat{A}\hat{\mathcal{U}}^\dagger)\mathcal{U}|\psi\rangle$$
$$= \langle\psi|\hat{A}|\psi\rangle$$

Thus all observable quantities (like inner products, eigen values, expectation values) remain unchanged under a unitary transformation.

Schrödinger picture
A particular choice of $\hat{\mathcal{U}}$ will give rise to a "particular approach" (called "picture"). We can choose $\hat{\mathcal{U}}$ to be the time evolution operator of the system: $\hat{\mathcal{U}} \equiv \hat{\mathcal{U}}(t, t_0) = \hat{U}(t, t_0)$. In the "picture" we already had in Chap. 3, Sect. 3.2, the state vectors evolve in time according to Eq. (3.13):

$$|\psi(t)\rangle = \hat{\mathcal{U}}(t, t_0)|\psi(t_0)\rangle = \hat{U}(t, t_0)|\psi(t_0)\rangle, \qquad \text{[Eq. (3.13)]}$$

where $\hat{\mathcal{U}}(t, t_0)$ is the time evolution operator (a unitary operator) and satisfies Eq. (3.19). An explicit expression was obtained and given by Eq. (3.20):

$$\hat{\mathcal{U}}(t, t_0) = \hat{U}(t, t_0) = \exp\Big(-\frac{i\hat{H}}{\hbar}(t - t_0)\Big) \qquad \text{[Eq. (3.20)]}.$$

This "picture" is usually referred to as the "Schrödinger picture". In this picture state vectors *evolve in time* according to Eq. (3.13), while operators are "stationary"; i.e. they *do not evolve in time*. They may have explicit time dependence (such that the operator is different at different times), but the operator itself does not "move" i.e. does not evolve in time.

Heisenberg picture

We can have an alternative approach, called Heisenberg picture, in which state vectors remain "stationary" (i.e. do not evolve in time), but operators "move" i.e. evolve in time (in addition to any explicit time dependence they may have). This we can do by taking $\hat{\mathcal{U}}$ to be the inverse time evolution operator $\hat{U}(t_0, t)$ from Eq. (3.20), so that $|\psi(t)\rangle$ becomes $|\psi(t_0)\rangle$ (independent of t):

$$\begin{aligned}\hat{\mathcal{U}}^{(H)} &= \hat{U}(t_0, t) = \hat{U}^{\dagger}(t, t_0) \\ &= \exp\Big(\frac{i\hat{H}}{\hbar}(t - t_0)\Big).\end{aligned} \tag{4.27}$$

To distinguish the pictures, we use superscripts (S) and (H) for quantities in Schrödinger and Heisenberg pictures respectively. Our earlier approach (in which state vectors "move") corresponds to Schrödinger picture and we rewrite the equation as

$$|\psi(t)\rangle^{(S)} = \hat{\mathcal{U}}^{(S)}|\psi(t_0)\rangle^{(S)} = \hat{U}(t, t_0)|\psi(t_0)\rangle^{(S)},$$

$$\text{with} \qquad \hat{\mathcal{U}}^{(S)} = \hat{U}(t, t_0) = \exp\Big(-\frac{i\hat{H}}{\hbar}(t - t_0)\Big). \tag{4.28}$$

The equation of motion satisfied by the state vector $|\psi(t)\rangle^{(S)}$ is the Schrödinger equation (hence this picture is called the Schrödinger picture)

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle^{(S)} = \hat{H}^{(S)}|\psi(t)\rangle^{(S)}. \tag{4.29}$$

An operator $\hat{A}^{(S)}$ in Schrödinger picture is "stationary" i.e.

$$\hat{A}^{(S)}(t) = \hat{A}^{(S)}(t_0) \equiv \hat{A}^{(S)}.$$

The state vector in the Heisenberg picture is

$$\begin{aligned}|\psi(t)\rangle^{(H)} &= \hat{\mathcal{U}}^{(H)}|\psi(t)\rangle^{(S)} = \hat{U}(t_0, t)|\psi(t)\rangle^{(S)} \\ &= \hat{U}(t_0, t)\hat{U}(t, t_0)|\psi(t_0)\rangle^{(S)} = \hat{U}(t_0, t_0)|\psi(t_0)\rangle^{(S)} \\ &= |\psi(t_0)\rangle^{(S)} \\ &= |\psi(t_0)\rangle^{(H)}.\end{aligned}$$

The last step follows from the previous step, which shows that $|\psi(t)\rangle^{(H)}$ is the same for all t (which is equal to $|\psi(t_0)\rangle^{(S)}$). Thus the state vectors in Heisenberg picture remain stationary. The state vectors in both the pictures coincide at the initial time t_0 only.

An operator in Heisenberg picture becomes, according to Eq. (4.26), where $\mathcal{U}$ is given by Eq. (4.27) together with Eq. (3.20)

$$\hat{A}^{(H)}(t) = \hat{\mathcal{U}}^{(H)}\hat{A}^{(S)}\hat{\mathcal{U}}^{(H)^\dagger} = \hat{U}^\dagger(t,t_0)\hat{A}^{(S)}\hat{U}(t,t_0)$$
$$= \exp\Big(\frac{i}{\hbar}\hat{H}(t-t_0)\Big)\hat{A}^{(S)}\exp\Big(-\frac{i}{\hbar}\hat{H}(t-t_0)\Big). \tag{4.30}$$

Since $\hat{U}(t_0,t_0) = \hat{1}$, we see that operators in both the pictures coincide at the initial time (t_0)

$$\hat{A}^{(H)}(t_0) = \hat{A}^{(S)}.$$

Also $\hat{H}$ commutes with $\hat{\mathcal{U}}^{(H)}$. Hence from Eq. (4.30)

$$\hat{H}^{(H)}(t) = \hat{H}^{(S)} \equiv \hat{H} \quad \text{at all times.}$$

Thus Hamiltonian is the same in both the pictures. Since in Heisenberg picture, only operators move, the equation of motion must be that of the operator. We can calculate the time derivative of Eq. (4.30), including a possible *explicit* time dependence of $\hat{A}^{(S)}$ as

$$\frac{\mathrm{d}\hat{A}^{(H)}(t)}{\mathrm{d}t} = \frac{\partial\hat{U}^\dagger(t,t_0)}{\partial t}\hat{A}^{(S)}\hat{U}(t,t_0) + \hat{U}^\dagger(t,t_0)\hat{A}^{(S)}\frac{\partial\hat{U}(t,t_0)}{\partial t}$$
$$+\hat{U}^\dagger(t,t_0)\frac{\partial\hat{A}^{(S)}}{\partial t}\hat{U}(t,t_0).$$

Since $\hat{U}(t,t_0)$ satisfies Eq. (3.19), we have (note that $\hat{H} = \hat{H}^{(S)}$)

$$\frac{\partial\hat{U}(t,t_0)}{\partial t} = \frac{1}{i\hbar}\hat{H}^{(S)}\hat{U}(t,t_0)$$

and taking Hermitian adjoint of this equation

$$\frac{\partial\hat{U}^\dagger(t,t_0)}{\partial t} = -\frac{1}{i\hbar}\hat{U}^\dagger(t,t_0)\hat{H}^{(S)}$$

Substituting these, we have

$$i\hbar\frac{\mathrm{d}\hat{A}}{\mathrm{d}t}^{(H)} = -\hat{U}^\dagger(t,t_0)\hat{H}^{(S)}\hat{A}^{(S)}\hat{U}(t,t_0) + \hat{U}^\dagger(t,t_0)\hat{A}^{(S)}\hat{H}^{(S)}\hat{U}(t,t_0)$$
$$+ i\hbar\hat{U}^\dagger(t,t_0)\frac{\partial\hat{A}^{(S)}}{\partial t}\hat{U}(t,t_0)$$
$$= -\hat{U}^\dagger(t,t_0)\hat{H}^{(S)}\hat{U}(t,t_0)\hat{U}^\dagger(t,t_0)\hat{A}^{(S)}\hat{U}(t,t_0)$$
$$+\hat{U}^\dagger(t,t_0)\hat{A}^{(S)}\hat{U}(t,t_0)\hat{U}^\dagger(t,t_0)\hat{H}^{(S)}\hat{U}(t,t_0) + i\hbar\hat{U}^\dagger(t,t_0)\frac{\partial\hat{A}^{(S)}}{\partial t}\hat{U}(t,t_0)$$
$$= \big[\hat{A}^{(H)},\hat{H}^{(H)}\big] + i\hbar\Big(\frac{\partial\hat{A}}{\partial t}\Big)^{(H)}, \tag{4.31}$$

using the definition of operators in Heisenberg picture, Eq. (4.30). Equation (4.31) is known as the Heisenberg equation of motion. If $\hat{A}^{(S)} \equiv \hat{A}$ has no explicit time dependence, the last term vanishes in Eq. (4.31).

In Heisenberg picture the operators change with time, which resembles (in a crude sense) the time dependence of classical dynamical variables. In Schrödinger picture, the time dependence of an observable is given by that of its expectation value, which is the same in both the pictures.

To visualize these two pictures, let us have a crude analogy. Suppose a train moves along a track in a field. To an observer in the field, the "state" of the train "moves", but the field remains "stationary". But for an observer in the train, his "state" remains "stationary", while the field "moves" backwards. The description of the observer in the field corresponds to the "Schrödinger picture", in which the "state" (train) evolves in time, while the "operator" (field) remains stationary. Any time-dependent change of the field (say a tractor tilling the field) corresponds to an explicit time dependence of the operator. The description of the observer in the train corresponds to the Heisenberg picture. His "state" (train) remains stationary, while the "operator" (field) evolves in time. In addition to this time evolution, he can also perceive the "explicit time dependence" (tractor tilling) of the "operator" (field). Both the pictures describe the same physical motion and are completely equivalent.

To obtain the time evolution of a particular case, we can follow either of the pictures. In Schrödinger picture, one has to solve the time-dependent Schrödinger equation, while in Heisenberg picture, one has to solve the Heisenberg equation of motion. For simple cases (see problems at the end of this Chapter), the Heisenberg equation of motion is easy to solve. For more complicated cases, the Schrödinger picture is more convenient and is commonly used.

Interaction picture

If the time-dependent Hamiltonian can be split into two parts

$$\hat{H} = \hat{H}_0 + \hat{V}(t),$$

such that the larger part (i.e. the part which has larger expectation value) $\hat{H}_0$ is independent of time and a smaller part $\hat{V}(t)$ depends on time, then it is useful to define an intermediate picture, called *interaction picture* [and indicated by a superscript (I)], in which time dependence of state vectors and operators are defined through

$$|\psi(t)\rangle^{(I)} = \exp\left(\frac{i\hat{H}_0}{\hbar}(t - t_0)\right)|\psi(t)\rangle^{(S)},$$

$$\hat{A}^{(I)}(t) = \exp\left(\frac{i\hat{H}_0}{\hbar}(t - t_0)\right)\hat{A}^{(S)}\exp\left(-\frac{i\hat{H}_0}{\hbar}(t - t_0)\right).$$

Clearly, at the initial time $t = t_0$, the state vector and the operator coincide with those of the Schrödinger picture. We can easily verify that

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle^{(I)} = \hat{V}^{(I)}(t)|\psi(t)\rangle^{(I)}, \tag{4.32}$$

$$i\hbar\frac{\mathrm{d}\hat{A}^{(I)}(t)}{\mathrm{d}t} = [\hat{A}^{(I)}(t), \hat{H}_0] + i\hbar\frac{\partial\hat{A}^{(I)}(t)}{\partial t}. \tag{4.33}$$

These are the basic equations for time evolution in the interaction picture. Equation (4.32) defines for the time evolution of a state vector, while Eq. (4.33) controls the time evolution of operators in this picture. Thus in the interaction picture both state vectors and operators move and their guiding equations are similar to the Schrödinger equation and the Heisenberg equation of motion, although their motions are governed by different parts (*viz.*, $\hat{V}$ and $\hat{H}_0$ respectively) of the Hamiltonian. If $\hat{V} = 0$, then $|\psi(t)\rangle^{(I)}$ becomes independent of time and this picture coincides with Heisenberg picture. On the other hand, if $\hat{H}_0 = 0$, then $\hat{A}^{(I)}(t)$ does not depend on time (except for any explicit time dependence), while Eq. (4.32) becomes the Schrödinger equation . Thus in this case, the interaction picture coincides with the Schrödinger picture. The interaction picture is convenient in time-dependent problems, where time dependence of the Hamiltonian is a weak perturbation over a dominant time-independent part.

4.7 Problems

1. Two matrices $\underline{\underline{A}}$ and $\underline{\underline{B}}$ anticommute. Also $\underline{\underline{A}}^2 = \underline{\underline{1}}$ and $\underline{\underline{B}}^2 = \underline{\underline{1}}$. Show that $\mathrm{Tr}(\underline{\underline{A}}) = 0$ and $\mathrm{Tr}(\underline{\underline{B}}) = 0$.
2. Show that a constant shift in potential does not have any effect on the wave function and energy is shifted by the same constant amount.
3. Consider a quantum particle with finite energy moving in one dimension encounters a region where $V(x) = +\infty$. Show that its wave function must vanish in that region.
4. Show that the expectation value of momentum is real.
5. Use the time-dependent Schrödinger equation to calculate

$$\frac{\mathrm{d}\langle\hat{A}\rangle}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\int\psi^*(\vec{r},t)\hat{A}(\vec{r},t)\psi(\vec{r},t)\mathrm{d}^3r,$$

   where $\psi(\vec{r},t)$ is normalized. Show that this is the same as the expectation value of the Heisenberg equation of motion, Eq. (4.31).
6. Find the eigen values and corresponding eigen vectors of

$$\underline{\underline{A}} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

7. An orthonormal basis in a two-dimensional Hilbert space is $\{|\phi_1\rangle, |\phi_2\rangle\}$. The Hamiltonian of the system is given by

$$\hat{H} = a\,|\phi_1\rangle\langle\phi_1| + b\,|\phi_2\rangle\langle\phi_2| + c\,|\phi_1\rangle\langle\phi_2| + \mathrm{d}\,|\phi_2\rangle\langle\phi_1|.$$

   (a) Can all the constants a, b, c, d be complex? Is there any relation between these constants?
   (b) Calculate the energy eigen values and corresponding eigen vectors in matrix representation.

8. The normalized wave function at time $t = 0$ of a one-dimensional case is

$$\psi(x, t = 0) = \Big(\frac{1}{2\pi\sigma^2}\Big)^{\frac{1}{4}} \exp(-\frac{x^2}{4\sigma^2}).$$

   Calculate $\psi(x, t)$ and the probability density in the Schrödinger picture. Show that the wave packet spreads, while its center remains stationary as time progresses.
   [Hint: Use the following relations (Ref. Blinder 1968)

$$\frac{\partial^2}{\partial x^2}\Big(\frac{1}{\sqrt{\sigma^2}} \exp(-\frac{x^2}{4\sigma^2})\Big) = \frac{\partial}{\partial(\sigma^2)}\Big(\frac{1}{\sqrt{\sigma^2}} \exp(-\frac{x^2}{4\sigma^2})\Big),$$

$$\exp\Big(\alpha\frac{\partial}{\partial z}\Big) f(z) = f(z + \alpha).$$

   Verify these relations.]

9. Obtain the Heisenberg equation of motion (HEM) for $\hat{x}$ and $\hat{p}$ for a free particle in one-dimension ($\hat{H} = \frac{\hat{p}^2}{2m}$) and then solve.
10. Obtain HEM for $\hat{x}$ for a particle moving in a one-dimensional potential, which in coordinate space is $V(x) = \frac{1}{2}kx^2 + \beta x$. Indicate how the HEM for $\hat{x}$ can be solved.

References

Arfken, G.: Mathematical Methods for Physicists. Academic Press, New York (1966)
Blinder, S.M.: Am. J. Phys. **36**, 525 (1968)
Chattopadhyay, P. K.: Mathematical Physics. New Age International (P) Ltd., New Delhi (2006)

Chapter 5
General Uncertainty Relation



Abstract General uncertainty relations for two incompatible observables are derived from the quantization postulate. Next we obtain the minimum uncertainty product and the corresponding wave packet.

Keywords Incompatible observables · Uncertainty relation satisfied by them · Derivation of uncertainty relation · Minimum uncertainty wave packet

In Chap. 1 we discussed the introduction of Heisenberg's uncertainty relations to explain experiments like diffraction of light by double slit. It was stated that the product of uncertainties in the measured values of a pair of conjugate variables must be at least of the order of $\hbar$. In this chapter, we present a mathematical derivation of the uncertainty relation. For this the *uncertainty* in the measurement of a dynamical variable needs a quantitative definition. This is defined as the root-mean-square (r.m.s.) deviation for a very large number of measurements. According to the fundamental postulates of quantum mechanics (Chap. 3, Sect. 3.2), when the system is in a general state (i.e. not an eigen state of the dynamical variable) results of measurement of that variable for the *same state* of the system can be *any one of the eigen values*. Hence the average[1] of deviations from the mean value vanishes, but an r.m.s. deviation is possible. This is because any deviation from the mean can be either positive or negative. But a squared deviation is always positive and its non-vanishing mean value is the mean-squared value.

5.1 Derivation of Uncertainty Relation

A pair of conjugate dynamical variables is represented by a pair of non-commuting Hermitian operators, say $\hat{A}$ and $\hat{B}$. From Chap. 2, Sect. 3.2, we know that if $\hat{A}$ and $\hat{B}$ are two observables with $[\hat{A}, \hat{B}] = 0$, then the system can be in the simultaneous eigen ket, say $|\Psi_{i\alpha}\rangle$, corresponding to eigen values a_i and b_α of $\hat{A}$ and $\hat{B}$ respectively[2], i.e.

[1] For the meaning of 'average' in this connection, see the discussion following Eq. (3.3).

[2] In general, if the indices i and α are *independent*, they can be considered as two quantum numbers. Hence in that case the quantum system is at least two-dimensional (see Chap. 8). An

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5_5

$$\hat{A}|\Psi_{i\alpha}\rangle = a_i|\Psi_{i\alpha}\rangle,$$
$$\hat{B}|\Psi_{i\alpha}\rangle = b_\alpha|\Psi_{i\alpha}\rangle.$$

Then according to postulate 2, corollary 2b, the state of the system *will not* change due to the measurements of $\hat{A}$ and $\hat{B}$, one after the other. Such observables are called "compatible (or commuting) observables". On the other hand, if $[\hat{A}, \hat{B}] \neq 0$, then $\hat{A}$ and $\hat{B}$ cannot have a common simultaneous eigen ket. Hence according to the postulate, a measurement of $\hat{A}$ on the system (which is in a general state $|\psi\rangle$) yields a particular eigen value (say a_i) and the state of the system collapses to the corresponding eigen ket $|\psi_i\rangle$ of $\hat{A}$. Since the state $|\psi\rangle$ is *not* an eigen ket of $\hat{B}$, a measurement of $\hat{B}$ will yield an eigen value (say) b_α and as a consequence the state of the system will collapse to the *corresponding* eigen ket $|\phi_\alpha\rangle$ of $\hat{B}$. Since this is *different* from $|\psi_i\rangle$ (as there are no *simultaneous* eigen ket of $\hat{A}$ and $\hat{B}$), a subsequent measurement of $\hat{A}$ will yield, in general, a *different* eigen value of $\hat{A}$ (not a_i). The same is true for the observable $\hat{B}$. Thus a large number of measurements of $\hat{A}$ and $\hat{B}$ will yield different results each time. Hence there will be a "spread" or "variance" in the results of measurements. Such observables (for which $[\hat{A}, \hat{B}] \neq 0$) are called "incompatible observables".

We consider simultaneous measurements of two incompatible observables $\hat{A}$ and $\hat{B}$ for the system in the state $|\psi\rangle$. Note that each single measurement is for the state $|\psi\rangle$, and the expectation value $\langle\cdots\rangle$ is with respect to this state, i.e. $|\psi\rangle$ [see discussion following Eq. (3.3)]. A simultaneous measurement of both $\hat{A}$ and $\hat{B}$ (for the same identically prepared state $|\psi\rangle$) will give a spread of measured values of each. Since the deviations $(\hat{A} - \langle\hat{A}\rangle)$ from the mean $(= \langle\hat{A}\rangle)$ can be either positive or negative and can take all possible values, the average of deviations would vanish:

$$\langle(\hat{A} - \langle\hat{A}\rangle)\rangle = \langle\hat{A}\rangle - \langle\langle\hat{A}\rangle\rangle = \langle\hat{A}\rangle - \langle\hat{A}\rangle = 0.$$

But the average of the squared deviations will be a positive quantity and is used to define the uncertainty (ΔA) (also called "variance") as (see Ref. Schiff 1968)

$$\Delta A = \langle(\hat{A} - \langle\hat{A}\rangle)^2\rangle^{\frac{1}{2}} = \langle\Big(\hat{A}^2 - \hat{A}\langle\hat{A}\rangle - \langle\hat{A}\rangle\hat{A} + \langle\hat{A}\rangle^2\Big)\rangle^{\frac{1}{2}}$$
$$= \sqrt{\langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2}.$$

Similarly

$$\Delta B = \langle(\hat{B} - \langle\hat{B}\rangle)^2\rangle^{\frac{1}{2}}$$
$$= \sqrt{\langle\hat{B}^2\rangle - \langle\hat{B}\rangle^2}. \tag{5.1}$$

Thus the "uncertainty" is defined as the root-mean-square deviation. We will find a relation between ΔA and ΔB when

exception is a free particle in one dimension having specific values of both energy and momentum. In this case, energy and momentum are not independent.

$$[\hat{A}, \hat{B}] = i\hat{C}. \tag{5.2}$$

An $i = \sqrt{-1}$ is included in the right side, so that $\hat{C}$ is Hermitian (i.e. also an observable). Now for any arbitrary state $|\psi\rangle$, let

$$|\psi_1\rangle = (\hat{A} - \langle\hat{A}\rangle)|\psi\rangle, \tag{5.3}$$

and

$$|\psi_2\rangle = (\hat{B} - \langle\hat{B}\rangle)|\psi\rangle. \tag{5.4}$$

Then for the state $|\psi\rangle$

$$\begin{aligned}(\Delta A)^2 &= \langle\psi|(\hat{A} - \langle\hat{A}\rangle)^2|\psi\rangle \\ &= \langle\psi|(\hat{A} - \langle\hat{A}\rangle)(\hat{A} - \langle\hat{A}\rangle)|\psi\rangle \\ &= \langle\psi_1|\psi_1\rangle \end{aligned} \tag{5.5}$$

(Since $\hat{A}$ is Hermitian and hence $\langle\hat{A}\rangle$ is real, the operator $(\hat{A} - \langle\hat{A}\rangle)$ is also Hermitian). Thus ΔA is the norm of $|\psi_1\rangle$

$$\Delta A = \| \, |\psi_1\rangle \, \|. \tag{5.6}$$

Similarly

$$\Delta B = \| \, |\psi_2\rangle \, \|. \tag{5.7}$$

Then by Schwarz's inequality Eq. (2.6)

$$\begin{aligned}(\Delta A)(\Delta B) &= \| \, |\psi_1\rangle \, \| \, \| \, |\psi_2\rangle \, \| \geq |\langle\psi_1|\psi_2\rangle| \\ &\geq |\mathrm{Im}\Big(\langle\psi_1|\psi_2\rangle\Big)| \\ &= |\frac{1}{2i}(\langle\psi_1|\psi_2\rangle - \langle\psi_2|\psi_1\rangle)|, \end{aligned} \tag{5.8}$$

where $\mathrm{Im}(\cdots)$ denotes "imaginary part of" $(\cdots)$. Similarly, $\mathrm{Re}(\cdots)$ denotes "real part of" $(\cdots)$. Note that the equality in the first inequality (in the first line) holds *only if* $|\psi_1\rangle$ and $|\psi_2\rangle$ are proportional (i.e. if they are in the same "direction"). The equality in the second step holds *only if* $\mathrm{Re}(\langle\psi_1|\psi_2\rangle) = 0$. Using Eqs. (5.3) and (5.4)

$$\begin{aligned}\Delta A \Delta B &\geq \frac{1}{2}\Big|\frac{1}{i}\Big(\langle\psi|(\hat{A} - \langle\hat{A}\rangle)(\hat{B} - \langle\hat{B}\rangle)|\psi\rangle - \langle\psi|(\hat{B} - \langle\hat{B}\rangle)(\hat{A} - \langle\hat{A}\rangle)|\psi\rangle\Big)\Big| \\ &= \frac{1}{2}\Big|\frac{1}{i}\langle\psi|(\hat{A}\hat{B} - \hat{B}\hat{A})|\psi\rangle\Big| \\ &= \frac{1}{2}\Big|\frac{i}{i}\langle\psi|\hat{C}|\psi\rangle\Big|. \end{aligned} \tag{5.9}$$

Hence we have

$$\Delta A \Delta B \geq \frac{|\langle C \rangle|}{2}. \tag{5.10}$$

This is the famous uncertainty relation for two non-commuting observables, known as *Heisenberg uncertainty relation*. For two conjugate variables $\hat{q}_i$ and $\hat{p}_i$ we have according to postulate 3 (Chap. 3, Sect. 3.2)

$$[\hat{q}_i, \hat{p}_j] = i\hbar\delta_{ij}.$$

Hence according to Eq. (5.10)

$$\Delta q_i \Delta p_j \geq \frac{\hbar}{2}\delta_{ij}. \tag{5.11}$$

For the well known relations, we have

$$[\hat{x}, \hat{p}_x] = i\hbar \quad \Rightarrow \quad \Delta x \Delta p_x \geq \frac{\hbar}{2}. \tag{5.12}$$

Similarly for y- and z-components. Also since $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$ [which are the Cartesian components of orbital angular momentum operator (see Chap. 10, Sect. 10.1)] satisfy commutation relations

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z \quad (x, y, z \text{ cyclic}), \tag{5.13}$$

we have

$$\Delta L_x \Delta L_y \geq \frac{\hbar}{2}|\langle \hat{L}_z \rangle| \quad (x, y, z \text{ cyclic}). \tag{5.14}$$

Since ϕ (the azimuthal angle in spherical polar coordinates) and $\hat{L}_z$ are two conjugate variables (note that $\hat{L}_z \rightarrow i\hbar\frac{\partial}{\partial\phi}$, so that $[\phi, \hat{L}_z] = i\hbar$), we have

$$\Delta\phi \Delta L_z \geq \frac{\hbar}{2}. \tag{5.15}$$

Another uncertainty relation between t and $\hat{E}$ can be found from the time-energy relation. Using Eq. (3.8), *viz.* $\hat{E} \rightarrow i\hbar\frac{\partial}{\partial t}$, we have

$$[t, \hat{E}] = -i\hbar.$$

Then from Eq. (5.10)

$$\Delta t \Delta E \geq \frac{\hbar}{2}. \tag{5.16}$$

However note that t is not a dynamical variable. So this relation is to be taken as a special relation [see discussion after Eq. (3.8)].

5.2 Minimum Uncertainty Product

During the derivation of the uncertainty relation Eq. (5.10), we used the $\geq$ sign in Eq. (5.8) twice. As mentioned immediately below Eq. (5.8), the first equality holds if

$$|\psi_1\rangle = g|\psi_2\rangle \quad (g \text{ is a constant, complex in general}), \tag{5.17}$$

and the second equality holds if

$$\mathrm{Re}\langle\psi_1|\psi_2\rangle = 0. \tag{5.18}$$

If these equations hold, we will get in Eq. (5.10)

$$\Delta A \Delta B = \frac{1}{2}|\langle\hat{C}\rangle|. \tag{5.19}$$

This is called the *minimum uncertainty product*, and the state $|\psi\rangle$ for which Eq. (5.19) holds is called the *minimum uncertainty state* (or $\psi(x) = \langle x|\psi\rangle$ as the minimum uncertainty wave packet). To get a specific form of the minimum uncertainty wave packet, we specify to the position-momentum uncertainty for a one-dimensional wave packet in coordinate representation. Thus we take

$$\hat{A} = \hat{x} \to x,$$

and

$$\hat{B} = \hat{p} \to -i\hbar\frac{d}{dx}.$$

Let

$$\hat{\alpha} = \hat{x} - \langle\hat{x}\rangle \to x - \langle x\rangle$$
$$\hat{\beta} = \hat{p} - \langle\hat{p}\rangle \to -i\hbar\frac{d}{dx} - \langle p\rangle.$$

Then $|\psi_1\rangle = \hat{\alpha}|\psi\rangle$, $|\psi_2\rangle = \hat{\beta}|\psi\rangle$.
For the minimum uncertainty product, conditions (5.17) and (5.18) give

$$\hat{\alpha}|\psi\rangle = g\hat{\beta}|\psi\rangle \tag{5.20}$$

and

$$\langle\psi|(\hat{\alpha}\hat{\beta} + \hat{\beta}\hat{\alpha})|\psi\rangle = 0. \tag{5.21}$$

In terms of the wave functions, Eq. (5.20) gives

$$\Big(x - \langle x\rangle\Big)\psi(x) = g\Big(-i\hbar\frac{d}{dx} - \langle p\rangle\Big)\psi(x)$$

or,

$$\frac{d\psi(x)}{dx} = \Big[\frac{i}{g\hbar}(x - \langle x\rangle) + \frac{i\langle p\rangle}{\hbar}\Big]\psi(x)$$

Integration gives

$$\psi(x) = N\exp\Big[\frac{i}{2g\hbar}(x - \langle x\rangle)^2 + \frac{i\langle p\rangle x}{\hbar}\Big] \tag{5.22}$$

Substitution of Eq. (5.20) in Eq. (5.21) gives

$$(g^* + g)\langle\psi|(\hat{\beta})^2|\psi\rangle = 0.$$

Now $\langle\psi|(\hat{\beta})^2|\psi\rangle\rangle \geq 0$. (note that $\hat{\beta}$ is Hermitian), and it can be zero only if $\hat{\beta}|\psi\rangle = 0$, i.e. only if $|\psi\rangle$ is an eigen ket of momentum. For this case, Eq. (5.20) shows that $\hat{\alpha}|\psi\rangle = 0$; i.e. it is also an eigen ket of position. But this is impossible since $[\hat{x}, \hat{p}] \neq 0$ and $|\psi\rangle$ cannot be a simultaneous eigen ket of both position and momentum. Then $g + g^* = 0$. Hence g is purely imaginary. Furthermore from the expression of Eq. (5.22), we see that $\psi(x)$ would diverge for $|x| \to \infty$ if g is *positive* imaginary. Hence for the wave function $\psi(x)$ to be square integrable (or finite as $|x| \to \infty$), g must be negative imaginary

$$g = -i\gamma \quad (\gamma \text{ is real positive})$$

Then from Eq. (5.22)

$$\psi(x) = N\exp\Big[-\frac{1}{2\gamma\hbar}\Big(x - \langle x\rangle\Big)^2 + \frac{i\langle p\rangle x}{\hbar}\Big]. \tag{5.23}$$

The constants N and γ can be obtained from the conditions

$$\int_{-\infty}^{+\infty} |\psi(x)|^2 dx = 1 \quad \text{(normalization)}, \tag{5.24}$$

and

$$\begin{aligned}(\Delta x)^2 &= \langle(x - \langle x\rangle)^2\rangle \\ &= \int_{-\infty}^{+\infty} \psi^*(x)\Big(x - \langle x\rangle\Big)^2\psi(x)dx \quad \text{(packet width)}.\end{aligned} \tag{5.25}$$

Calculating N and γ from Eqs. (5.24) and (5.25) and substituting back in Eq. (5.23), we get the normalized minimum uncertainty wave packet (in one dimension) (see Ref. Schiff 1968)

$$\psi(x) = \left[2\pi(\Delta x)^2\right]^{-\frac{1}{4}} \exp\left[-\frac{(x-\langle x\rangle)^2}{4(\Delta x)^2} + \frac{i\langle p\rangle x}{\hbar}\right]. \tag{5.26}$$

We will see in Chap. 8, Sect. 8.5 that the corresponding time-dependent wave function $\psi(x,t)$ is a traveling wave packet (traveling in the positive x direction if $\langle p\rangle > 0$), which has a Gaussian envelope (i.e. a Gaussian profile) of width Δx. For this packet $\Delta x \Delta p = \frac{\hbar}{2}$, which is the minimum possible value for $\Delta x \Delta p$.

5.3 Problems

1. If $\hat{A}$ and $\hat{B}$ are Hermitian and $[\hat{A}, \hat{B}] = i\hat{C}$, then show that $\hat{C}$ is also Hermitian.
2. If $[\hat{H}, \hat{A}] = 0$ and $\hat{A}$ not explicitly dependent on time, show that (ΔA) does not change with time.
3. In Eq. (5.14), consider the state to be an eigen ket of angular momentum in standard representation $|\psi\rangle = |lm\rangle$ (see Chap. 10). Then for the state $|l, 0\rangle$, the product $\Delta L_x \Delta L_y$ can vanish (since $\langle L_z\rangle = m\hbar = 0$). But L_x and L_y do not commute and are not simultaneously measurable. Why is this so? Calculate ΔL_x and ΔL_y.

Reference

Schiff, L.I.: Quantum Mechanics, 3rd edn. McGraw-Hill Book Company Inc., Singapore (1968)

Chapter 6
Harmonic Oscillator: Operator Method



Abstract One-dimensional harmonic oscillator is used as an example of solving a quantum mechanical problem in the abstract Hilbert space, to obtain eigen values, eigen vectors and matrix elements. Connection with coordinate space representation is shown by obtaining the wave function.

Keywords Importance of harmonic oscillator in physics · Eigen solutions by ladder operator · Matrix elements · Coordinate space wave function from abstract states · Uncertainty relation

6.1 Importance of Simple Harmonic Oscillator

As an example of how one can solve a quantum mechanical problem in the abstract Hilbert space ($\mathcal{H}$) using abstract vectors and operators in $\mathcal{H}$, we consider a simple harmonic oscillator (SHO) in one dimension.

The SHO is one of the most important problems in QM. This is a simple example to illustrate basic concepts and methods. Also from the point of view of physics, SHO is particularly useful, since for any physical system, whether simple or complex, the equilibrium position corresponds to the minimum of the potential energy in the classical sense. At the minimum of any potential function of one or more variables, the first derivative (or the gradient in multi-dimension) must vanish and the second derivative must be positive (for multi-dimension the matrix of the second derivatives must be positive definite), so that the equilibrium position is a local minimum of the potential function (or surface). Hence a Taylor series expansion about the minimum always leads to, for small deviations, a potential having a quadratic dependence on the deviation from the equilibrium position. Since a constant shift in potential is unimportant in non-relativistic treatment, this corresponds to a SHO for small deviations from equilibrium position. This is the reason we find applications of SHO in almost all branches of physics – molecular, solid state, nuclear, quantum field theory, etc. Historically the discrete energies of radiation oscillator (an SHO) proposed by Planck led to the birth of quantum mechanics. Although in reality such potentials have finite extent, but for simplicity in mathematical idealization, we assume the SHO potential to extend to infinity.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5_6

Since SHO is a good approximation to the potential near the equilibrium point, the potential can be approximated to be a sum of an SHO plus a perturbation. The fact that eigen solutions of the SHO are known analytically (see below) makes this procedure very convenient for the perturbation treatment (to be discussed in Volume 2). However, note that any realistic potential *cannot exactly* be taken as a sum of an ideal SHO plus a perturbation, since an ideal SHO becomes infinite at large separations. On the other hand, this very fact makes its eigen functions decrease very rapidly as the deviation from the equilibrium point increases, leading to rapid convergence of matrix elements of the perturbation. For this reason, the CONS of eigen functions of SHO is a convenient basis for the expansion of a trial wave function in a variational treatment (to be discussed in Volume 2). Thus the study of SHO is of great interest in quantum mechanics (in fact for all of physics).

6.2 Energy Eigen Values and Eigen Vectors

We follow Dirac's elegant operator method. This, indeed, is the operator method in the abstract Hilbert space. For simplicity, we consider a particle moving in a one-dimensional, time-independent harmonic oscillator potential. The time-independent Hamilton operator is

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2, \tag{6.1}$$

where $\omega = \sqrt{\frac{K}{m}}$ is the classical angular frequency, K being the stiffness constant (so that potential is $V(x) = \frac{1}{2}Kx^2$) and m is mass of the particle. Note that $\hat{x}$, $\hat{p}$ are Hermitian and satisfy

$$[\hat{x}, \hat{p}] = i\hbar. \tag{6.2}$$

Introduce a non-Hermitian operator ($\hat{a}$) and its Hermitian adjoint

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + \frac{i\hat{p}}{m\omega}\right),$$
$$\hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} - \frac{i\hat{p}}{m\omega}\right). \tag{6.3}$$

The operators $\hat{a}$ and $\hat{a}^\dagger$ will be called *annihilation* and *creation* operators for reasons to be clear later. Now

$$
\begin{aligned}
[\hat{a}, \hat{a}^\dagger] &= \left(\frac{m\omega}{2\hbar}\right)\left[\hat{x} + \frac{i\hat{p}}{m\omega}, \hat{x} - \frac{i\hat{p}}{m\omega}\right] \\
&= \left(\frac{m\omega}{2\hbar}\right)\left(-\frac{i}{m\omega}[\hat{x}, \hat{p}] + \frac{i}{m\omega}[\hat{p}, \hat{x}]\right) \\
&= \left(\frac{m\omega}{2\hbar}\right)\left(-\frac{i}{m\omega}(i\hbar) + \frac{i}{m\omega}(-i\hbar)\right) \\
&= 1. \qquad (6.4)
\end{aligned}
$$

Next define another operator

$$\hat{N} = \hat{a}^\dagger \hat{a}. \qquad (6.5)$$

Note that $\hat{N}$ is hermitian. We will see later that $\hat{N}$ is the number operator. From Definition 6.3, we have

$$
\begin{aligned}
\hat{N} &= \hat{a}^\dagger \hat{a} \\
&= \left(\frac{m\omega}{2\hbar}\right)\left(\hat{x}^2 + \frac{\hat{p}^2}{m^2\omega^2} - \frac{i}{m\omega}(\hat{p}\hat{x} - \hat{x}\hat{p})\right) \\
&= \left(\frac{m\omega}{2\hbar}\right)\left(\frac{\hat{p}^2}{m^2\omega^2} + \hat{x}^2 - \frac{\hbar}{m\omega}\right) \\
&= \frac{1}{\hbar\omega}\left(\frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2\right) - \frac{1}{2} \\
&= \frac{\hat{H}}{\hbar\omega} - \frac{1}{2}. \qquad (6.6)
\end{aligned}
$$

Hence

$$\hat{H} = \hbar\omega\left(\hat{N} + \frac{1}{2}\right), \qquad (6.7)$$

and $[\hat{H}, \hat{N}] = 0$. Therefore, both can be simultaneously specified, i.e. both can have simultaneous eigen kets. Consider the eigen value equation of $\hat{N}$

$$\hat{N}|n\rangle = n|n\rangle, \qquad (6.8)$$

where n is the eigen value corresponding to the normalized eigen ket $|n\rangle$. Then from Eqs. (6.7) and (6.8), we have

$$
\begin{aligned}
\hat{H}|n\rangle &= \hbar\omega\left(\hat{N} + \frac{1}{2}\right)|n\rangle \\
&= \hbar\omega\left(n + \frac{1}{2}\right)|n\rangle. \qquad (6.9)
\end{aligned}
$$

Thus $|n\rangle$ is also an eigen ket of $\hat{H}$ corresponding to the eigen value $E_n = (n + \frac{1}{2})\hbar\omega$.

$$\begin{aligned} \hat{H}|n\rangle &= E_n|n\rangle \\ &\Rightarrow E_n = \left(n + \tfrac{1}{2}\right)\hbar\omega. \end{aligned} \tag{6.10}$$

To understand the physical significance of $\hat{a}$ and $\hat{a}^\dagger$, let us calculate

$$\begin{aligned} [\hat{N}, \hat{a}] &= [\hat{a}^\dagger\hat{a}, \hat{a}] \\ &= [\hat{a}^\dagger, \hat{a}]\hat{a} + \hat{a}^\dagger[\hat{a}, \hat{a}] \\ &= -\hat{a} + 0 \\ &= -\hat{a}. \end{aligned} \tag{6.11}$$

Again

$$\begin{aligned} [\hat{N}, \hat{a}^\dagger] &= [\hat{a}^\dagger\hat{a}, \hat{a}^\dagger] \\ &= [\hat{a}^\dagger, \hat{a}^\dagger]\hat{a} + \hat{a}^\dagger[\hat{a}, \hat{a}^\dagger] \\ &= \hat{a}^\dagger. \end{aligned} \tag{6.12}$$

Then

$$\begin{aligned} \hat{N}\hat{a}^\dagger|n\rangle &= ([\hat{N}, \hat{a}^\dagger] + \hat{a}^\dagger\hat{N})|n\rangle \\ &= (\hat{a}^\dagger + \hat{a}^\dagger\hat{N})|n\rangle \\ &= (n + 1)\hat{a}^\dagger|n\rangle. \end{aligned} \tag{6.13}$$

Thus $\hat{a}^\dagger|n\rangle$ is an eigen ket of $\hat{N}$ corresponding to eigen value $(n + 1)$; hence by Eq. (6.8), it must be proportional to $|n + 1\rangle$ (since there is no degeneracy in one dimension, see Sect. 8.8.3)

$$\hat{a}^\dagger|n\rangle = c_+|n + 1\rangle, \tag{6.14}$$

where c_+ is a constant, complex in general. Take inner product of Eq. (6.14) with itself

$$\langle n|\hat{a}.\hat{a}^\dagger|n\rangle = |c_+|^2\langle n + 1|n + 1\rangle = |c_+|^2, \tag{6.15}$$

assuming that all $|n\rangle$ are normalized. Then left hand side (LHS) of Eq. (6.15) is

$$\begin{aligned} \text{LHS} &= \langle n|[\hat{a}, \hat{a}^\dagger] + \hat{a}^\dagger\hat{a}|n\rangle \\ &= \langle n|(\hat{1} + \hat{N})|n\rangle \\ &= n + 1. \end{aligned} \tag{6.16}$$

Hence by Eqs. (6.15) and (6.16)

$$c_+ = \sqrt{n+1}. \tag{6.17}$$

(Taking the arbitrary phase to be zero). Hence from Eq. (6.14)

$$\hat{a}^\dagger |n\rangle = \sqrt{n+1}\ |n+1\rangle. \tag{6.18}$$

In a similar manner, we have

$$\begin{aligned} \hat{N}\hat{a}|n\rangle &= ([\hat{N}, \hat{a}] + \hat{a}\hat{N})|n\rangle \\ &= (-\hat{a} + \hat{a}\hat{N})|n\rangle \\ &= (n-1)\hat{a}|n\rangle. \end{aligned} \tag{6.19}$$

Thus $\hat{a}|n\rangle$ is an eigen ket of $\hat{N}$ corresponding to eigen value $(n-1)$, hence by Eq. (6.8) it is proportional to $|n-1\rangle$

$$\hat{a}|n\rangle = c_-|n-1\rangle, \tag{6.20}$$

c_- being another constant. Take inner product of Eq. (6.20) with itself

$$\begin{aligned} \langle n|\hat{a}^\dagger\hat{a}|n\rangle &= |c_-|^2\langle n-1|n-1\rangle = |c_-|^2 \\ &= \langle n|\hat{N}|n\rangle \\ &= n \end{aligned} \tag{6.21}$$

(assuming $|n\rangle$ to be normalized). Hence

$$|c_-|^2 = n, \tag{6.22}$$

or

$$c_- = \sqrt{n}. \tag{6.23}$$

(Again taking the arbitrary phase to be zero). Hence

$$\hat{a}|n\rangle = \sqrt{n}\ |n-1\rangle. \tag{6.24}$$

Applying $\hat{a}$ repeatedly on Eq. (6.24), we have

$$\begin{aligned} (\hat{a})^2|n\rangle &= \hat{a}\sqrt{n}\ |n-1\rangle = \sqrt{n}\sqrt{n-1}\ |n-2\rangle, \\ (\hat{a})^3|n\rangle &= \hat{a}\sqrt{n(n-1)}\ |n-2\rangle = \sqrt{n(n-1)(n-2)}\ |n-3\rangle, \\ &\text{and so on.} \end{aligned} \tag{6.25}$$

Thus we get a new eigen ket of $\hat{N}$ with eigen value reduced by one, each time we apply $\hat{a}$ on it. Now the norm of a vector $|\psi\rangle = \hat{a}|n\rangle$ is

$$\langle\psi|\psi\rangle = \langle n|\hat{a}^\dagger\hat{a}|n\rangle = \langle n|\hat{N}|n\rangle = n. \tag{6.26}$$

Therefore, $n \geq 0$ and

$$\begin{aligned} n = 0 \quad \text{iff} \quad |\psi\rangle &= 0, \qquad \text{the null vector.} \\ \text{Thus we have} \qquad \hat{a}|0\rangle &= 0, \qquad \text{the null vector.} \end{aligned} \tag{6.27}$$

By applying $\hat{a}$ repeatedly, the eigen value n of $|n\rangle$ can be reduced indefinitely using Eq. (6.25), unless the factor in front vanishes at some stage. It can vanish only if n is a non-negative integer. Then applying $\hat{a}$ operator n-times on $|n\rangle$, we have

$$\begin{aligned} (\hat{a})^n|n\rangle &= \sqrt{n(n-1)(n-2)\cdots 3\cdot 2\cdot 1}\;|0\rangle \\ &= \sqrt{n!}\;|0\rangle. \end{aligned} \tag{6.28}$$

Applying $\hat{a}$ again and using Eq. (6.27)

$$\hat{a}(\hat{a})^n\;|n\rangle = \sqrt{n!}\;\hat{a}\;|0\rangle \;\propto \sqrt{n!}\times 0 = 0, \tag{6.29}$$

using Eq. (6.27). The right most side of Eq. (6.29) becomes the null vector. Applying $\hat{a}$ any number of times on this, results in the null vector again. Thus if n is a non-negative integer, the norm of Eq. (6.26) can never become negative. Hence, n can take *only non-negative integer values*

$$n = 0, 1, 2, 3, \cdots, \tag{6.30}$$

and

$$\hat{a}|0\rangle = 0. \tag{6.31}$$

Then from Eq. (6.10),

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \qquad (n = 0, 1, 2, \cdots) \tag{6.32}$$

gives the energy eigen values. For the ground state (i.e. lowest energy state) $n = 0$ and $E_0 = \frac{1}{2}\hbar\omega$. We can now apply $\hat{a}^\dagger$ successively on $|0\rangle$ (using Eq. (6.18)) to get the exited states. Putting $n = 0, 1, 2\cdots$ in Eq. (6.18))

$$\begin{aligned} n = 0 \to |1\rangle &= \frac{1}{\sqrt{1}}\hat{a}^\dagger\,|0\rangle \\ n = 1 \to |2\rangle &= \frac{1}{\sqrt{2}}\hat{a}^\dagger\,|1\rangle = \frac{1}{\sqrt{1.2}}(\hat{a}^\dagger)^2\,|0\rangle \\ n = 2 \to |3\rangle &= \frac{1}{\sqrt{3}}\hat{a}^\dagger\,|2\rangle = \frac{1}{\sqrt{3!}}(\hat{a}^\dagger)^3\,|0\rangle, \end{aligned} \tag{6.33}$$

and in general

$$|n\rangle = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}|0\rangle. \tag{6.34}$$

We can easily verify that each of $|n\rangle$ is normalized if $|0\rangle$ is. The orthonormality of these states is given by

$$\langle n'|n\rangle = \delta_{nn'} \tag{6.35}$$

The number n is the excitation quantum number and $\hat{N}$ counts the number of excitations – hence called number operator. The operator $\hat{a}^\dagger$ creates one extra unit of excitation, according to Eq. (6.18) – hence it is called the "creation") operator. These are also called *ladder operators*. On the other hand Eq. (6.24) shows that $\hat{a}$ destroys (or annihilates) one unit of excitatio – hence it is called the "destruction" (or "annihilation") operator. The representation in which $\hat{N}$ is diagonal is called the *occupation number representation*, value of n being the quantum state occupied.

Thus we can solve the one-dimensional harmonic oscillator problem to get the eigen values and eigen kets using operators in the abstract vector space. In this method we can calculate all observables as seen below. Finally we can also get the wave function in coordinate or momentum representation from the eigen kets.

6.3 Matrix Elements

We can easily calculate the matrix elements of $\hat{x}$ and $\hat{p}$ in this occupation number representation. From Eq. (6.3) we have

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a} + \hat{a}^\dagger) \tag{6.36}$$

$$\hat{p} = i\sqrt{\frac{m\hbar\omega}{2}}(\hat{a}^\dagger - \hat{a}) \tag{6.37}$$

Then using Eqs. (6.18) and (6.24),

$$\begin{aligned}\langle n'|\hat{x}|n\rangle &= \sqrt{\frac{\hbar}{2m\omega}}(\langle n'|\hat{a}|n\rangle + \langle n'|\hat{a}^\dagger|n\rangle) \\ &= \sqrt{\frac{\hbar}{2m\omega}}(\sqrt{n}\langle n'|n-1\rangle + \sqrt{n+1}\langle n'|n+1\rangle) \\ &= \sqrt{\frac{\hbar}{2m\omega}}(\sqrt{n}\,\delta_{n',n-1} + \sqrt{n+1}\,\delta_{n',n+1}),\end{aligned} \tag{6.38}$$

and similarly

$$\langle n'|\hat{p}|n\rangle = i\sqrt{\frac{m\hbar\omega}{2}}(\sqrt{n+1}\,\delta_{n',n+1} - \sqrt{n}\,\delta_{n',n-1}). \tag{6.39}$$

Matrix elements of higher powers of $\hat{x}$ and $\hat{p}$ can be obtained by introducing the completeness relation

$$\sum_m |m\rangle\langle m| = \hat{1},$$

and then using already known matrix element of lower powers. For example, matrix element of $\hat{x}^2$ can be obtained by

$$\langle n'|\hat{x}^2|n\rangle = \sum_m \langle n'|\hat{x}|m\rangle\langle m|\hat{x}|n\rangle,$$

and using Eq. (6.38) for each of the matrix elements in the above. Similarly, matrix element of $\hat{x}^3$ can be obtained in terms of matrix elements of $\hat{x}$ and $\hat{x}^2$, and so on. Matrix elements of all powers of $\hat{p}$ can likewise be obtained in terms of Eq. (6.39).

Note that both matrices $\underline{\underline{x}}$ and $\underline{\underline{p}}$ are Hermitian. In this representation Hamiltonian is a diagonal matrix

$$\langle n'|\hat{H}|n\rangle = E_n\,\delta_{n',n} = \left(n + \frac{1}{2}\right)\hbar\omega\,\delta_{n'n}. \tag{6.40}$$

Thus we get Heisenberg's matrix mechanics approach in which position, momentum and the Hamiltonian are represented by Hermitian matrices.

6.4 Coordinate Space Wave Function

We can also calculate the wave functions in the coordinate space representation, by taking coordinate projection of the abstract vector equations (6.18) and (6.24). For the ground state, we take inner product of Eq. (6.27) with $\langle x|$ (eigen bra of position operator, corresponding to eigen value x)

$$\langle x|\hat{a}|0\rangle = 0, \tag{6.41}$$

or, using Eq. (6.3)

$$\langle x|\left(\sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + \frac{i\,\hat{p}}{m\omega}\right)\right)|0\rangle = 0.$$

Introducing the completeness relation between an operator and $|0\rangle$ and then using the appropriate form of the *local* operator, we have

$$\langle x|\hat{x}|0\rangle = \int \langle x|\hat{x}|x'\rangle \mathrm{d}x' \langle x'|0\rangle = \int \mathrm{d}x' \delta(x-x') x' \langle x'|0\rangle = x\langle x|0\rangle,$$

and

$$\langle x|\hat{p}|0\rangle = \int \langle x|\hat{p}|x'\rangle \mathrm{d}x' \langle x'|0\rangle = \int \mathrm{d}x' \delta(x-x')(-i\hbar\frac{\mathrm{d}}{\mathrm{d}x'})\langle x'|0\rangle = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\langle x|0\rangle.$$

substituting these, we have

$$\sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{i}{m\omega}\left(-i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\right)\right)\langle x|0\rangle = 0.$$

Writing $\langle x|0\rangle = \psi_0(x)$ (the coordinate space wave function for the ground state), we have

$$\left(x + \frac{\hbar}{m\omega}\frac{\mathrm{d}}{\mathrm{d}x}\right)\psi_0(x) = 0.$$

Let us define $x_0 = \sqrt{\frac{\hbar}{m\omega}}$, which has the dimension of length and sets the length scale of the oscillator. Thus

$$\frac{\mathrm{d}\psi_0(x)}{\mathrm{d}x} = -\frac{x}{x_0^2}\psi_0(x).$$

This can be easily integrated to get

$$\psi_0(x) = A\,\exp\Big(-\frac{1}{2}\left(\frac{x}{x_0}\right)^2\Big),$$

where A is an arbitrary constant. To get it, we normalize ψ_0 according to

$$\begin{aligned}
1 &= \int_{-\infty}^{+\infty} |\psi_0(x)|^2 \mathrm{d}x \\
&= |A|^2 \int_{-\infty}^{+\infty} \mathrm{e}^{-(\frac{x}{x_0})^2}\mathrm{d}x \\
&= |A|^2 x_0 \int_{-\infty}^{+\infty} \mathrm{e}^{-y^2}\mathrm{d}y \\
&= |A|^2 x_0\sqrt{\pi} \qquad\qquad \Rightarrow\ A = \frac{1}{\pi^{\frac{1}{4}}\sqrt{x_0}},
\end{aligned}$$

and

$$\psi_0(x) = \frac{1}{\pi^{\frac{1}{4}}\sqrt{x_0}} \exp\Big(-\frac{1}{2}\left(\frac{x}{x_0}\right)^2\Big), \qquad \Big(x_0 = \sqrt{\frac{\hbar}{m\omega}}\Big). \tag{6.42}$$

We can obtain the coordinate space wave function of the nth exited state by taking the coordinate projection of Eq. (6.34) and using Eqs. (6.3) and (6.42)

$$\begin{aligned}
\psi_n(x) &= \langle x|n\rangle = \langle x|\frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}|0\rangle \\
&= \frac{\langle x|}{\sqrt{n!}}\Big(\sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} - \frac{i\hat{p}}{m\omega}\right)\Big)^n|0\rangle \\
&= \frac{1}{\sqrt{n!}}\Big(\frac{1}{\sqrt{2}x_0}\Big)^n\Big(x - x_0^2\frac{\mathrm{d}}{\mathrm{d}x}\Big)^n\langle x|0\rangle \\
&= \Big(\frac{1}{\sqrt{\pi}2^n n!}\Big)^{\frac{1}{2}}\Big(\frac{1}{x_0^{n+\frac{1}{2}}}\Big)\Big(x - x_0^2\frac{d}{dx}\Big)^n \exp\Big(-\frac{1}{2}\left(\frac{x}{x_0}\right)^2\Big)
\end{aligned} \tag{6.43}$$

This can be directly evaluated for a specific $n = 1, 2, \cdots$. The wave function for a general nth excited state can be put in the more familiar form, using the standard recurrence relations of Hermite polynomials (see Arfken 1966; Chattopadhyay 2006). The result is (see Chap. 9, Sect. 9.4)

$$\psi_n(x) = \left(\frac{1}{x_0\sqrt{\pi}2^n n!}\right)^{\frac{1}{2}} \mathrm{H}_n\left(\frac{x}{x_0}\right)\exp\left(-\frac{1}{2}\left(\frac{x}{x_0}\right)^2\right), \tag{6.44}$$

where $\mathrm{H}_n(\frac{x}{x_0})$ is the Hermite polynomial of degree n. Thus we get complete solutions in coordinate space by the operator method.

6.5 Uncertainty Relation

We can also calculate the expectation values $\langle\hat{x}^2\rangle$, $\langle\hat{p}^2\rangle$ and the uncertainty for position and momentum. Using Eq. (6.36)

$$\hat{x}^2 = \frac{\hbar}{2m\omega}(\hat{a}^2 + (\hat{a}^\dagger)^2 + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}). \tag{6.45}$$

For the ground state, we take matrix element of Eq. (6.45) between $\langle 0|$ and $|0\rangle$. Then since

$$\begin{aligned}
&\hat{a}|0\rangle = 0 \\
\text{and} \qquad &\langle 0|\hat{a}^\dagger = 0,
\end{aligned} \tag{6.46}$$

only the third term of Eq. (6.45) contributes

$$\begin{aligned}\langle 0|\hat{x}^2|0\rangle &= \frac{\hbar}{2m\omega}\langle 0|\hat{a}\hat{a}^\dagger|0\rangle \\ &= \frac{\hbar}{2m\omega}\langle 0|([\hat{a},\hat{a}^\dagger]+\hat{a}^\dagger\hat{a})|0\rangle \\ &= \frac{\hbar}{2m\omega}\langle 0|(1+\hat{a}^\dagger\hat{a})|0\rangle \\ &= \frac{\hbar}{2m\omega} \\ &= \frac{x_0^2}{2}. \end{aligned} \tag{6.47}$$

Similarly

$$\begin{aligned}\langle 0|\hat{p}^2|0\rangle &= \frac{m\hbar\omega}{2} \\ &= \frac{\hbar^2}{2}\left(\frac{m\omega}{\hbar}\right) \\ &= \frac{\hbar^2}{2x_0^2}. \end{aligned} \tag{6.48}$$

Now from Eqs. (6.38) and (6.39)

$$\langle 0|\hat{x}|0\rangle = 0, \qquad \langle 0|\hat{p}|0\rangle = 0. \tag{6.49}$$

Then the uncertainties for the ground state are

$$(\Delta x)_0^2 = \langle 0|\hat{x}^2|0\rangle - \langle 0|\hat{x}|0\rangle^2 = \frac{x_0^2}{2}, \tag{6.50}$$

and

$$(\Delta p)_0^2 = \langle 0|\hat{p}^2|0\rangle - \langle 0|\hat{p}|0\rangle^2 = \frac{\hbar^2}{2x_0^2}. \tag{6.51}$$

Hence

$$(\Delta x)_0^2(\Delta p)_0^2 = \frac{\hbar^2}{4} \qquad \Rightarrow \qquad (\Delta x)_0(\Delta p)_0 = \frac{\hbar}{2}. \tag{6.52}$$

Thus for the ground state of the oscillator, we get the minimum uncertainty product. This is understandable, since the ground state wave function is Gaussian (as we saw earlier in Chap. 5, Sect. 5.2 that the minimum uncertainty wave function has to have a Gaussian profile). For the nth exited state, we can show that

$$(\Delta x)_n(\Delta p)_n = \left(n + \frac{1}{2}\right)\hbar. \tag{6.53}$$

6.6 Problems

1. Prove both Eqs. (6.34) and (6.35) by mathematical induction.
2. Prove Eq. (6.44) from Eq. (6.43), using the standard recurrence relations of Hermite polynomials (see Arfken 1966; Chattopadhyay 2006). *Hint:* Use mathematical induction method.
3. Calculate $\langle n'|\hat{x}^2|n\rangle$ and $\langle n'|\hat{p}^2|n\rangle$ for the one-dimensional harmonic oscillator, using the operator method.
4. Using the previous result, prove Eq. (6.53).
5. Calculate $\langle n'|\hat{x}^3|n\rangle$ and $\langle n'|\hat{p}^3|n\rangle$ using Eqs. (6.38) and (6.39).
6. Prove that $(\Delta x)_n(\Delta p)_n = (n + \frac{1}{2})\hbar$.

References

Arfken, G.: Mathematical Methods for Physicists. Academic Press, New York (1966)
Chattopadhyay, P.K.: Mathematical Physics. New Age International (P) Ltd., New Delhi (2006)

Chapter 7
Mathematical Preliminary II: Theory of Second Order Differential Equations

Abstract For the benefit of students not familiar with theory of second order differential equations, it is provided together with theory of boundary value problems in this chapter. We demonstrate the connection between physics and mathematics. A number of standard differential equations, their standard solutions and properties in brief have been included. These are useful and provide handy references for solving quantum mechanical problems.

Keywords Singularities of second order differential equations · Frobenius method · Boundary value problem · Sturm-Liouville theory · Mathematics-Physics connection · Standard differential equations

From now on, we will follow the standard presentation of Quantum Mechanics in the position coordinate representation we discussed in Chap 4, Sect. 4.1. As a first step, for simplicity, we tackle the problem of a non-relativistic single particle of mass m moving in a time-independent and velocity-independent potential $V(\vec{r})$ in three-dimensional space. Then we have to solve Eq. (4.4) with $V(\vec{r}, t) = V(\vec{r})$

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r}, t) = \Big[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})\Big]\psi(\vec{r}, t).$$

Since V is time-independent we can easily separate time by substituting

$$\psi(\vec{r}, t) = u(\vec{r})T(t).$$

Substitution and division of both sides by $\psi = uT$ result in

$$i\hbar\frac{1}{T(t)}\frac{\mathrm{d}T(t)}{\mathrm{d}t} = \frac{1}{u(\vec{r})}\Big[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})\Big]u(\vec{r}) = E.$$

In the last equation, since left side is a function of t only and the middle is a function of $\vec{r}$ only, each side must be a constant, *independent of* t *and* $\vec{r}$, which

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5_7

we call E (separation constant). This results in two separate differential equations: integration of the t equation gives

$$T(t) = A \exp\left(-\frac{i}{\hbar} Et\right), \tag{7.1}$$

where A is a constant, which we absorb in the normalization constant of $u(\vec{r})$. Equating the middle to E, we get the *time-independent Schrödinger equation*

$$\Big[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})\Big]u(\vec{r}) = Eu(\vec{r}). \tag{7.2}$$

Equation (7.2) is the eigen value equation for the Hamiltonian $\hat{H} = \big[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})\big]$, with eigen value E. Hence E must be real and is interpreted as the energy of the system (in general, all separation constants represent physical quantities). For a three-dimensional case, Eq. (7.2) is a partial differential equation.

The wave function $\psi(\vec{r}, t)$ is given by

$$\psi(\vec{r}, t) = u(\vec{r}) \exp\left(-\frac{i}{\hbar} Et\right),$$

and $u(\vec{r})$ is normalized as (since E is real)

$$\int |\psi(\vec{r}, t)|^2 \mathrm{d}^3 r = \int |u(\vec{r})|^2 \mathrm{d}^3 r = 1,$$

where the integral is over all space.

In many cases, we face a one-dimensional (or effectively one-dimensional) motion. In one dimension (with variable x) we replace Eq. (7.2) by

$$\Big[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\Big]u(x) = Eu(x). \tag{7.3}$$

This is a second order differential equation in one variable. Its solution, subject to appropriate boundary conditions, provides the energy E and the wave function $u(x)$. Since Eq. (7.2) is linear and homogeneous in u, the latter can be known up to a proportionality constant, called *normalization constant*. It is obtained from the standard normalization condition

$$\int |u(x)|^2 \mathrm{d}x = 1. \tag{7.4}$$

The lower and upper limits of integration are such as to cover the entire range (*domain*) of x appropriate for the particle moving in the potential $V(x)$.

The time-dependent wave function becomes (overall normalization constant is included in $u(x)$)

$$\psi(x,t) = u(x)T(t) = u(x)\mathrm{e}^{-\frac{i}{\hbar}Et}.$$

Hence the probability density at position x at time t is

$$P(x,t) = \left|\psi(x,t)\right|^2 = \left|u(x)\right|^2 \quad (\text{since } E \text{ is real}).$$

Thus the position probability density becomes independent of t. Hence the state represented by $u(x)$ is called a stationary state and time-independent Schrödinger equation is called the *stationary state Schrödinger equation*.

For a spherically symmetric potential, such that $V(\vec{r})$ does not depend on the polar angles, i.e. $V(\vec{r}) = V(r)$, Eq. (7.2) reduces to a differential equation in one variable (r), which resembles the one-dimensional Schrödinger equation. However note that this is not really a one-dimensional Schrödinger equation, since there are contributions in this equation from eigen values arising from other variables. For example, the eigen value of (θ, ϕ)-dependent equation gives rise to an additional centrifugal repulsion term $\frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}$ to $V(r)$. Also the interval is not that of a one-dimensional motion. See Sect. 10.3, Chap. 10, as also the footnote of Sect. 7.1.5 of this chapter.

Before seeking solutions of Eq. (7.3) for a specified potential, subject to normalization (7.4), we will discuss in the following the theory of such equations.

7.1 Second Order Differential Equations

In this section, we study some general properties of the solutions of a second order linear homogeneous differential equation like Eq. (7.3). Note that depending on the potential $V(\vec{r})$ and the choice of coordinate system, Eq. (7.2) may or may not be separable in the variables. Even if separable in a particular case, the resulting equation may not be as simple as Eq. (7.3). In general, it will involve first derivatives also. Besides quantum mechanics, such equations appear in many other branches of physics. So we consider a general form

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2} + P(x)\frac{\mathrm{d}}{\mathrm{d}x} + Q(x)\right]y(x) = 0, \tag{7.5}$$

or in short (a prime denotes a differentiation of a function of x once w.r.t. the argument x)

$$y''(x) + P(x)y'(x) + Q(x)y(x) = 0,$$

where x is in the domain $a \le x \le b$ (the *interval* of x is denoted by $[a, b]$, where a and b are called *boundaries*) and $P(x)$ and $Q(x)$ are two analytic functions of x in this interval. This means that both $P(x)$ and $Q(x)$ have no singularities in the open interval $a < x < b$. There may be singularities at one or both the boundaries $x = a$

and $x = b$. Eq. (7.3) is a special case, with $P(x) = 0$ and $Q(x) = \frac{2m}{\hbar^2}[E - V(x)]$. The restrictions on $P(x)$ and $Q(x)$ are usually necessary for problems in physics and we will discuss more on these in Sect. 7.1.4. In the following we discuss some general properties.

7.1.1 Singularities of the Differential Equation

A differential equation can be classified according to its singularities at one or more points within the interval $[a, b]$. Any point x_0 is said to be an ordinary point or a singular point according to the following definitions.

Def. 1. Ordinary point: A point x_0 in the interval is said to be an ordinary point if both $P(x)$ and $Q(x)$ are analytic at $x = x_0$.

Def. 2. Singular point: A point x_0 is said to be a singular point if either

$$\lim_{x \to x_0} P(x)$$

diverges or

$$\lim_{x \to x_0} Q(x)$$

diverges or both diverge.

There are two types of singularities: *regular* (or *removable*) and *irregular* (or *essential*) singularity according to the following definitions.

Def. 3. regular singular point: If x_0 is a singular point and both

$$\lim_{x \to x_0} (x - x_0) P(x)$$

and

$$\lim_{x \to x_0} (x - x_0)^2 Q(x)$$

are finite, then x_0 is said to be a regular singular point of the differential equation.

Def. 4. Irregular singular point:

On the other hand if either

$$\lim_{x \to x_0} (x - x_0) P(x)$$

diverges or

$$\lim_{x \to x_0} (x - x_0)^2 Q(x)$$

diverges or both diverge, then x_0 is an irregular singular point of the differential equation.

The nature of singularity as $x \to \infty$ can be found by changing the variable x to $z = \frac{1}{x}$ in Eq. (7.5) and studying the resulting equation in z for its singularity as $z \to 0$.

It is expected that the analytic behavior of the solution $y(x)$ at $x = x_0$ will depend on the nature of singularity of the differential equation at that point. As the analyticity of the solution is of particular interest for the physics problem, the knowledge of singularities of a differential equation is very important.

7.1.2 Linear Dependence of the Solutions

We discussed linear dependence and independence of vectors in Chap. 2, Sect. 2.1.3. We also noted there that a set of functions can be regarded as vectors. So we can expect similar definitions for a set of functions, say, the set of solutions of Eq. (7.5). First we note a few important properties of the solutions. Proofs of most of them are straightforward.

Property 1:
Since Eq. (7.5) is linear and homogeneous in its solution, if $y_1(x)$ is a solution, then $cy_1(x)$, where c is a constant (independent of x) is also a solution. Hence the solution of Eq. (7.5) is *normalizable* up to a multiplying constant.

Property 2:
If $y_1(x)$ and $y_2(x)$ are two different solutions of Eq. (7.5), then any linear combination $c_1 y_1(x) + c_2 y_2(x)$, where c_1 and c_2 are constants, is also a solution.

Def. 1. Linear dependence of a set of functions:
Consider a set of n functions $\{\phi_1(x), \phi_2(x), \cdots, \phi_n(x)\}$. This set of functions is linearly dependent if there exists a relation

$$c_1\phi_1(x) + c_2\phi_2(x) + \cdots + c_n\phi_n(x) = 0,$$

such that *not all the n constants* $\{c_1, c_2, \cdots, c_n\}$ *vanish* (at least two should be non-vanishing).

Def. 2. Linear independence of a set of functions:
The set of functions $\{\phi_1(x), \phi_2(x), \cdots, \phi_n(x)\}$ is linearly independent if there exists no relation of the type

$$c_1\phi_1(x) + c_2\phi_2(x) + \cdots + c_n\phi_n(x) = 0,$$

where $\{c_1, c_2, \cdots, c_n\}$ are n constants, except the *trivial solution* $c_1 = c_2 = \cdots = c_n = 0$.

Test for linear dependence or independence of a set of functions:
Consider a set of n analytic (i.e. differentiable any number of times) functions $\{\phi_1(x), \phi_2(x), \cdots, \phi_n(x)\}$ and a set of n constants $\{c_1, c_2, \cdots, c_n\}$ satisfying the equation

$$c_1\phi_1(x) + c_2\phi_2(x) + \cdots + c_n\phi_n(x) = 0. \tag{7.6}$$

Differentiating Eq. (7.6) w.r.t. x once, twice, $\cdots$, $(n-1)$ times, we have

$$\begin{aligned} c_1\phi_1'(x) + c_2\phi_2'(x) + \cdots + c_n\phi_n'(x) &= 0 \\ c_1\phi_1''(x) + c_2\phi_2''(x) + \cdots + c_n\phi_n''(x) &= 0 \\ &\cdots \\ c_1\phi_1^{(n-1)}(x) + c_2\phi_2^{(n-1)}(x) + \cdots + c_n\phi_n^{(n-1)}(x) &= 0, \end{aligned} \tag{7.7}$$

where $\phi_i^{(m)}(x)$ means the mth derivative of $\phi_i(x)$. The set of n Eqs. (7.6) and (7.7) can be considered as a set of n linear homogeneous equations for the n unknowns $\{c_1, c_2, \cdots, c_n\}$, for a given set of n functions $\{\phi_1(x), \phi_2(x), \cdots, \phi_n(x)\}$. Let $W(x)$ be the determinant formed by the coefficients of $\{c_1, c_2, \cdots, c_n\}$

$$W(x) = \begin{vmatrix} \phi_1(x) & \phi_2(x) & \cdots & \phi_n(x) \\ \phi_1'(x) & \phi_2'(x) & \cdots & \phi_n'(x) \\ \cdots & \cdots & \cdots & \cdots \\ \phi_1^{(n-1)}(x) & \phi_2^{(n-1)}(x) & \cdots & \phi_n^{(n-1)}(x) \end{vmatrix}. \tag{7.8}$$

This determinant $W(x)$ is called the *Wronskian* of the set of functions $\{\phi_1(x), \phi_2(x), \cdots, \phi_n(x)\}$. The set of linear homogeneous equations has a non-trivial solution for the unknown constants *only if the determinant of their coefficient* $W(x)$ *vanishes*. Then by Def. 1, the set of functions is linearly dependent. On the other hand, if $W(x) \neq 0$, then the only solution for the linear homogeneous set of Eqs. (7.6) and (7.7) is the trivial solution $c_1 = c_2 = \cdots = c_n = 0$. Hence the set of functions is linearly independent according to Def. 2.

Thus to test the linear dependence of a set of functions $\{\phi_1(x), \phi_2(x), \cdots, \phi_n(x)\}$, we have to construct the determinant $W(x)$ according to Eq. (7.8) and

$$\begin{aligned} &\text{if } W(x) = 0, \text{ then the set is linearly dependent,} \\ &\text{if } W(x) \neq 0, \text{ then the set is linearly independent.} \end{aligned} \tag{7.9}$$

Note that the Wronskian has to vanish identically at all points of the interval for the set of functions to be linearly dependent. If the Wronskian is zero at isolated points, it does not guarantee that the set of functions is linearly dependent, since at values of x for which $W(x) \neq 0$ only the trivial solution $c_1 = c_2 = \cdots = c_n = 0$ is possible (these coefficients are *constants* and have the same value for all x), whereas the trivial solution is always possible even at points where $W(x) = 0$.

Property 3:
Equation (7.5) can have *at most two linearly independent solutions.*

Proof:
Let $y_1(x)$, $y_2(x)$ and $y_3(x)$ be three solutions of Eq. (7.5). The Wronskian of these functions is

$$\begin{aligned} W(x) &= \begin{vmatrix} y_1(x) & y_2(x) & y_3(x) \\ y_1'(x) & y_2'(x) & y_3'(x) \\ y_1''(x) & y_2''(x) & y_3''(x) \end{vmatrix} \\ &= \begin{vmatrix} y_1 & y_2 & y_3 \\ y_1' & y_2' & y_3' \\ (-Py_1' - Qy_1) & (-Py_2' - Qy_2) & (-Py_3' - Qy_3) \end{vmatrix} \\ &= 0. \end{aligned}$$

In the above we used Eq. (7.5) to go from the first line to the second line. (for shortness we leave out the argument x of functions in the second line). This determinant vanishes, since the third row is $(-P)$ times the second row plus $(-Q)$ times the first row. This is true at all values of x. Hence the Wronskian vanishes identically and by Eq. (7.9), the set $\{y_1(x), y_2(x), y_3(x)\}$ is linearly dependent. On the other hand if we consider *only two* solutions $y_1(x)$ and $y_2(x)$, the 2×2 determinant (involving *only the first derivatives*) does not vanish automatically (unless $y_1(x)$ and $y_2(x)$ are proportional), and hence they are linearly independent. So we can find at most two linearly independent solutions of Eq. (7.5).

Also note that if there are two solutions $y_1(x)$ and $y_2(x)$, which *are not proportional*, then they must be linearly independent.

7.1.3 Series Solution: Frobenius Method

Here we discuss a method of solution of the second order differential equation, which is generally applicable to all such equations for the physical problem. For a physically meaningful solution, the wave function must be *well behaved*, which means that the wave function must be (besides being square-integrable) analytic everywhere or at least piecewise continuous everywhere, so that its first derivative may have discontinuities at isolated points (where the potential function has infinite discontinuities. See Sect. 8.8.1). Hence, at most, only one of the two linearly independent solutions may have possible singularities, which can be discarded in favor of the other solution. Then the differential equation can have singularities (usually only at one or both the boundaries), which may be reflected in the singularity of its solution. Thus the wave function must be analytic in the entire open interval $a < x < b$ or its sub-intervals (i.e. piecewise analytic). Consequently a series expansion of the solution about any point in the open interval or sub-interval is possible. This approach is called the

Frobenius method. In this method, a singular infinite series may be converted to a 'well behaved' solution by imposing a condition on the parameters of the equation, which truncates the infinite series to a polynomial. This condition (in general the condition leading to well behaved solution) gives the eigen value.

For the solution of Eq. (7.5), we substitute

$$y(x) = \sum_{\lambda=0}^{\infty} a_\lambda (x - x_0)^{k+\lambda} \quad (a_0 \neq 0) \tag{7.10}$$

for an expansion about x_0 in the open interval $a < x < b$. The restriction $a_0 \neq 0$ is imposed in order to fix a value for the still undetermined k, called the *index* (see Problems). The index k and the coefficients $a_0, a_1, a_2, \cdots$ will be obtained by substituting Eq. (7.10) in Eq. (7.5). We will demonstrate the method by a specific example in the following.

Example: Classical harmonic oscillator equation
As a simple example, we consider the classical harmonic oscillator equation

$$y''(x) + \omega^2 y(x) = 0, \tag{7.11}$$

where $y(x)$ represents the displacement at x and the constant ω is the frequency of the classical oscillator. Since $P(x) = 0$ and $Q(x) = \omega^2$ are both finite in $-\infty < x < \infty$, this equation has no singularities for finite x. Putting $z = \frac{1}{x}$, we can see that this equation has irregular singularities at $z = 0$ (i.e. as $x \to \pm\infty$). Consequences of these will become clear later. For a series solution about $x = 0$, we substitute Eq. (7.10) with $x_0 = 0$ in Eq. (7.11) (note that since the solution is analytic, the infinite sum converges and differentiation can be done term by term) we have

$$\sum_{\lambda=0}^{\infty} a_\lambda (k + \lambda)(k + \lambda - 1) x^{k+\lambda-2} + \omega^2 \sum_{\lambda=0}^{\infty} a_\lambda x^{k+\lambda} = 0. \tag{7.12}$$

Left side of Eq. (7.12) is a power series. since a power series expansion is unique, coefficient of each power in x must vanish. Equating the coefficient of the *lowest* power of x to zero will give an equation for the unknown index k. It is called the *indicial equation*. The lowest power of x is $(k - 2)$, coming from the first term only, for $\lambda = 0$. Hence

$$\text{Coefficient of } x^{(k-2)} = a_0 k(k - 1) = 0.$$

Since $a_0 \neq 0$, we have

$$k(k - 1) = 0. \tag{7.13}$$

This is the indicial equation. This is a quadratic equation in k and the solutions are

$$k = 1 \quad \text{and} \quad k = 0$$

For each value of k, Eq. (7.10) provides a solution. They are not proportional, hence are linearly independent. Thus $k = 1$ and $k = 0$ will give respectively the *first* and the *second* solution. We will follow the convention that the numerically larger index will be referred to as the first solution.

Some general properties:
We note that for a second order differential equation (as will result from the Schrödinger equation), the lowest power of k will (in general) come from the second derivative. Consequently, the indicial equation will in general be a quadratic equation in k. The two roots will in general give *two linearly independent* solutions (except when the roots are identical or differ by an integer for expansion about a regular singular point, see later). For expansion about an ordinary point, the roots are always $k = 1$ and $k = 0$ (see below). A general linear combination of the two linearly independent solutions is just appropriate as a general solution, as the *boundary conditions* at the two boundaries can be imposed by choosing the two arbitrary constants appropriately.

Coming back to Eq. (7.12), equating coefficients of higher powers of x will give the expansion coefficients $a_1, a_2, \cdots$. The next power gives

$$\text{Coefficient of } x^{(k-1)} = a_1 k(k+1) = 0. \tag{7.14}$$

For $k = 1$, we must have $a_1 = 0$. But for $k = 0$, a_1 need not vanish. Next we can have a *recurrence relation* for a_i with $i \geq 2$ by setting the coefficient of algebraically the largest general power of x in Eq. (7.12) to zero

$$\text{Coefficient of } x^{(k+j)} = a_{j+2}(k+j+2)(k+j+1) + \omega^2 a_j = 0, \qquad (j \geq 0)$$

which gives

$$a_{j+2} = -\frac{\omega^2}{(k+j+2)(k+j+1)}\, a_j. \qquad (j \geq 0). \tag{7.15}$$

Setting $j = 0, 2, 4, \cdots$ we get a_2, a_4, a_6 $\cdots$ in terms of a_0, a_2, a_4 $\cdots$ respectively, and using the already known lower coefficients, we can get *all even* coefficients in terms of a_0 only. Similarly we get *all odd* coefficients in terms of a_1 only. The solutions involve two arbitrary coefficients. Thus apparently we get two linearly independent solutions for $k = 0$ and one more (with a different arbitrary a_0) for $k = 1$ [for which a_1 must vanish, see Eq. (7.14)]. But we know at most two linearly independent solutions are possible. We will see below that we can choose $a_1 = 0$, so that all odd coefficients vanish by Eq. (7.15) and we get two linearly independent solutions corresponding to $k = 1$ and $k = 0$. We can verify that for the choice $a_1 \neq 0$ (only possible for $k = 0$), the two solutions thus obtained will be: one proportional to the $k = 1$ solution and the other a linear combination of this one and the one with choice $k = 0$ and $a_1 = 0$. We leave its proof as an exercise (see Problems). Since a_1 is *not related* to a_0 by Eq. (7.15) in this example, we can choose $a_1 = 0$ *without loss of generality*. Then by Eq. (7.15), all odd coefficients vanish.

$$\text{Choose } a_1 = 0, \text{ then } a_i = 0 \quad (i = 3, 5, 7, \cdots).$$

First solution, $k = 1$:
Putting $k = 1$ in Eq. (7.15)

$$a_{j+2} = -\frac{\omega^2}{(j+3)(j+2)}\, a_j. \qquad (j \geq 0)$$

Putting $j = 0, 2, 4, \cdots$ we have

$$\begin{aligned} j = 0 \;\Rightarrow\; a_2 &= -\frac{\omega^2}{3 \times 2} a_0 = -\frac{\omega^2}{3!} a_0 \\ j = 2 \;\Rightarrow\; a_4 &= -\frac{\omega^2}{5 \times 4} a_2 = (-1)^2 \frac{\omega^4}{5!} a_0 \\ j = 4 \;\Rightarrow\; a_6 &= -\frac{\omega^2}{7 \times 6} a_4 = (-1)^3 \frac{\omega^6}{7!} a_0 \\ &\cdots . \end{aligned}$$

By inspection, we can identify the general term (which can be proved easily by mathematical induction)

$$a_{2n} = (-1)^n \frac{\omega^{2n}}{(2n+1)!} a_0 \quad (n = 0, 1, 2, \cdots)$$

Hence we have the first solution

$$\begin{aligned} y_1(x) = y(x)|_{k=1} &= \sum_{\lambda=0}^{\infty} a_\lambda x^{k+\lambda} = \sum_{n=0}^{\infty} a_{2n} x^{1+2n} \\ &= \frac{a_0}{\omega} \sum_{n=0}^{\infty} (-1)^n \frac{(\omega x)^{2n+1}}{(2n+1)!} = \frac{a_0}{\omega} \sin(\omega x) \end{aligned} \qquad (7.16)$$

This is a well known solution. The constant a_0 is arbitrary and $\frac{a_0}{\omega}$ can be taken as the normalization constant.

Second solution, $k = 0$
Putting $k = 0$ in Eq. (7.15)

$$a_{j+2} = -\frac{\omega^2}{(j+2)(j+1)} a_j \quad (j = 0, 1, 2, \cdots) \qquad (7.17)$$

As we discussed earlier, for $k = 0$, Eq. (7.14) is already satisfied and there is no need for a_1 to vanish. But from Eq. (7.17), we see that a_1 is independent of a_0, and there is no loss of generality if we choose $a_1 = 0$. Then

$$\text{choose } a_1 = 0, \quad \Rightarrow \quad a_3 = a_5 = \cdots = 0.$$

Thus the odd series vanishes. We can see that choosing $a_1 \neq 0$, the odd series gives rise to a solution proportional to the first solution (see Problems). Thus we do not lose generality by the choice $a_1 = 0$. Hence the second solution is obtained from Eq. (7.17) with $j = 0, 2, 4, \cdots$ (the even series):

$$\begin{aligned} j = 0 \Rightarrow \quad a_2 &= -\frac{\omega^2}{2 \times 1} a_0 = -\frac{\omega^2}{2!} a_0 \\ j = 2 \Rightarrow \quad a_4 &= -\frac{\omega^2}{4 \times 3} a_2 = (-1)^2 \frac{\omega^4}{4!} a_0 \\ j = 4 \Rightarrow \quad a_6 &= -\frac{\omega^2}{6 \times 5} a_4 = (-1)^3 \frac{\omega^6}{6!} a_0 \\ &\cdots \end{aligned}$$

In general

$$a_{2n} = (-1)^n \frac{\omega^{2n}}{(2n)!} a_0 \quad (n = 0, 1, 2, \cdots)$$

Thus the second solution is

$$\begin{aligned} y_2(x) = y(x)|_{k=0} &= \sum_{\lambda=0}^{\infty} a_\lambda x^{k+\lambda} = \sum_{n=0}^{\infty} a_{2n} x^{0+2n} \\ &= a_0 \sum_{n=0}^{\infty} (-1)^n \frac{(\omega x)^{2n}}{(2n)!} = \alpha_0 \cos(\omega x) \end{aligned} \tag{7.18}$$

The undetermined constant a_0 is the new normalization constant, unrelated to a_0 of the first solution. The first and second solutions are the well known solution of Eq. (7.11).

General nature of series solution

In the above example, we get both linearly independent solutions by the series method, for a series expansion about $x = 0$. Note that $x = 0$ is an ordinary point of the differential equation. We already discussed that if the series expansion is about a regular or an irregular singular point, the series solution is likely to show the singular nature of the solution. Exploring a variety of second order differential equations (see Problems) we can arrive at the following observations:

- If the expansion is about an ordinary point, we get both solutions by series method.
- If the expansion is about a regular singular point, at least one solution in series form is possible.
- If the expansion is about an irregular singular point, we may or may not get any solution by series method.

In problems of physics, all points in the domain of physical interest are either ordinary points or at worst may have a few isolated regular singular points. This indeed is to be expected, as the solution (in order that it is physically meaningful) has to be analytic (hence expandable in a series) over the entire physical domain, or at worst, in piecewise segments of it. Hence, considering $x = 0$ as the point of expansion, $P(x)$ and $Q(x)$ of Eq. (7.5) can have, at worst, first order and second order singularities respectively at $x = 0$. Writing $P(x) = \frac{p(x)}{x}$ and $Q(x) = \frac{q(x)}{x^2}$, such that $p(x)$ and $q(x)$ are analytic at $x = 0$, the most general form, with $x = 0$ being at worst a regular singular point, becomes (multiplying through by x^2)

$$x^2 y''(x) + x p(x) y'(x) + q(x) y(x) = 0.$$

Since $p(x)$ and $q(x)$ are analytic at $x = 0$, they can be expanded in power series, starting with zero power

$$p(x) = \sum_{m=0}^{\infty} p_m x^m \quad \text{and} \quad q(x) = \sum_{m=0}^{\infty} q_m x^m.$$

If $x = 0$ is an ordinary point, we must have $p_0 = q_0 = q_1 = 0$. Substituting a series solution of the form of Eq. (7.10) with $x_0 = 0$, we can arrive at the following conclusions (see Bell 1968)

1. If $x = 0$ is an ordinary point then $k_1 = 1$ and $k_2 = 0$ and both solutions can be found in series form.
2. One series solution, corresponding to the first (larger, k_1) root of the indicial equation, can always be found.
3. If the difference of roots $(k_1 - k_2)$ is not an integer or zero, both series solutions can be found.
4. If $k_1 - k_2$ is an integer, the second solution in a series form breaks down in general. An exception is the case when $k_1 = 1$ and $k_2 = 0$, when $x = 0$ is an ordinary point and both series solutions are possible.
5. If $k_1 = k_2$, we get only one linearly independent solution in series form.

From the above, we can conclude that *at least one* solution in series form can always be found, if the expansion is about an ordinary point or at worst a regular singular point.

7.1.4 Boundary Value Problem: Sturm–Liouville Theory

The fundamental equation of quantum mechanics is the Schrödinger equation [Eq. (3.12) in coordinate representation]. This is an equation which gives non-trivial solutions only for special values, i.e. *eigen values* for the energy (E), as we will see in Chap. 8. Such an equation is called an *eigen value equation*. Indeed postulate

2 states that any observable, represented by a Hermitian operator must satisfy an eigen value equation (3.1). The Schrödinger equation results when the operator is the Hamiltonian. In coordinate representation this equation becomes a differential eigen value equation. The solution must satisfy physical (in the present case, quantum mechanical) acceptability conditions, which restrict the constant E to a set of special values called eigen values. Corresponding solutions are called eigen functions.

We discussed that all Hermitian operators representing observables can be in several forms – as an abstract operator in the Hilbert space, as a differential operator (as in coordinate and momentum representations) or as a Hermitian matrix. The operator can also be represented in an integral operator form, which will not be discussed here. The time-independent Schrödinger equation is given as an abstract operator eigen value equation by Eq. (3.11). The same is given as a differential eigen value equation by Eq. (3.12) and as a matrix eigen value equation by Eq. (4.22). In Chap. 6, we saw an example of the abstract operator equation. But the most common applications of quantum mechanics involve coordinate representation. Eq. (3.12) leads to a second order linear homogeneous differential equation in each separated variable. Each one will then have two linearly independent solutions, involving two arbitrary constants. Now the solution (wave function) must satisfy boundary conditions at the two boundaries of the interval $[a, b]$. These boundary conditions determine the arbitrary constants. Such cases are referred to as *boundary value problems*, where the condition of well behavior of the solution (quantum acceptability) is satisfied only for special allowed values (*eigen values*) for E.

An important requirement in quantum mechanics is that the operator representing an observable must be Hermitian. In the following we see how it is defined for the differential operator.

Consider a general second order differential operator in the interval $a \leq x \leq b$

$$\hat{\mathcal{L}}(x) = p(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + r(x)\frac{\mathrm{d}}{\mathrm{d}x} + q(x),$$

such that

$$\hat{\mathcal{L}}(x)u(x) = p(x)u''(x) + r(x)u'(x) + q(x)u(x), \tag{7.19}$$

subject to the conditions

1. $p(x)$, $q(x)$ and $r(x)$ are real and continuous in $a \leq x \leq b$.
2. $p'(x)$, $p''(x)$ and $r'(x)$ are continuous in $a \leq x \leq b$.
3. $p(x)$ has no zeros in the interior of the interval $a < x < b$, which means that the equation

$$\hat{\mathcal{L}}(x)u(x) = 0$$

has no singularities in $a < x < b$. However, $p(x)$ may vanish at either one or both the boundaries. Then the differential equation can have singularities at one or both boundaries. This happens quite often in physical cases.

We next define an adjoint operator

$$\hat{\tilde{\mathcal{L}}}(x)\{ \;=\; \frac{\mathrm{d}^2}{\mathrm{d}x^2}\{p(x) - \frac{\mathrm{d}}{\mathrm{d}x}\{r(x) + q(x)\{,$$

such that

$$\begin{aligned}\hat{\tilde{\mathcal{L}}}(x)u(x) &= \frac{\mathrm{d}^2}{\mathrm{d}x^2}\{p(x)u(x)\} - \frac{\mathrm{d}}{\mathrm{d}x}\{r(x)u(x)\} + q(x)\{u(x)\} \\ &= p(x)u''(x) + \{2p'(x) - r(x)\}u'(x) \\ &\quad +\{p''(x) - r'(x) + q(x)\}u(x). \end{aligned} \tag{7.20}$$

Def. Self adjoint operator:
The operator $\hat{\mathcal{L}}(x)$ is said to be *self-adjoint*, if

$$\hat{\tilde{\mathcal{L}}}(x) = \hat{\mathcal{L}}(x).$$

For this to be true, we have from Eqs. (7.19) and (7.20)

$$2p'(x) - r(x) = r(x),$$

and

$$p''(x) - r'(x) + q(x) = q(x).$$

Both these are satisfied if

$$p'(x) = r(x).$$

If this condition is satisfied, we have from Eq. (7.19)

$$\begin{aligned}\hat{\mathcal{L}}(x) &= p(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + p'(x)\frac{\mathrm{d}}{\mathrm{d}x} + q(x) \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\Big[p(x)\frac{\mathrm{d}}{\mathrm{d}x}\Big] + q(x). \end{aligned} \tag{7.21}$$

This form of $\hat{\mathcal{L}}(x)$ is self-adjoint and is called the Sturm–Liouville (SL) form. We can easily verify that any homogeneous second order differential equation

$$\hat{\mathcal{L}}(x)u(x) = p(x)u''(x) + r(x)u'(x) + q(x)u(x) = 0$$

can be put in SL form by pre-multiplying it with a function (see Problems)

$$f(x) = \frac{1}{p(x)}\exp\Big(\int \frac{r(x)}{p(x)}\mathrm{d}x\Big).$$

According to postulates of quantum mechanics, an operator representing an observable has to be Hermitian. A self-adjoint abstract operator ($\hat{A}^\dagger = \hat{A}$) or a self-adjoint matrix ($\underline{\underline{A}}^\dagger = \underline{\underline{A}}$) is Hermitian. However a second order differential operator in SL form (self-adjoint) is not automatically Hermitian. This is because its solutions are not unique until boundary conditions are imposed. In this connection we first discuss a second order differential eigen value equation.

Second order differential eigen value equation:
We have seen that a second order linear homogeneous differential equation can always be put in the SL form by multiplying it with a suitable function $f(x)$. An eigen value equation satisfied by $\hat{\mathcal{L}}(x)$ in SL form is

$$\hat{\mathcal{L}}(x)u_\lambda(x) + \lambda w(x)u_\lambda(x) = 0, \tag{7.22}$$

where $\hat{\mathcal{L}}(x)$ is in SL form given by Eq. (7.21). Here λ is a constant which can take only special allowed values, called *eigen values* and the corresponding solution $u_\lambda(x)$, subject to boundary conditions (see below), is called the *eigen function* belonging to eigen value λ. The function $w(x)$ [which is the final result after multiplying the original equation by $f(x)$] is called the *weight function*. It is a real and positive function in the interval $a \le x \le b$, except that it may have isolated zeros within this interval.

As mentioned above, we need additional boundary conditions for $\hat{\mathcal{L}}(x)$ to be Hermitian. We first see how a differential Hermitian operator is defined. Later, we will discuss certain properties (usually referred to as theorems) satisfied by eigen value equations of Hermitian operators.

The form of differential eigen value equation (7.22) usually appears in mathematics texts, while the eigen value term is kept on the right side for physical problems.

Def. Differential Hermitian operator:
Consider a second order differential operator $\hat{\mathcal{L}}$ in SL form, Eq. (7.21), in the interval $a \le x \le b$. This operator is said to be Hermitian with respect to functions $u(x)$ and $v(x)$ (which may be complex in general) if

$$\int_a^b v^*(x)\hat{\mathcal{L}}(x)u(x)\mathrm{d}x = \int_a^b u(x)\hat{\mathcal{L}}(x)v^*(x)\mathrm{d}x. \tag{7.23}$$

(See the end of this subsection for justification of this condition). Using Eq. (7.21) for $\hat{\mathcal{L}}$ and integrating the left side of Eq. (7.23) by parts twice, we can put the left side of Eq. (7.23) as the right side plus two surface terms. Hence these extra surface terms must vanish. This gives

$$v^*(x)p(x)u'(x)|_{x=a} = v^*(x)p(x)u'(x)|_{x=b}, \tag{7.24}$$

and

$$u(x)p(x)v^{*\prime}(x)|_{x=a} = u(x)p(x)v^{*\prime}(x)|_{x=b}. \tag{7.25}$$

These are the conditions that must be satisfied in order that the differential operator $\hat{\mathcal{L}}$ given by Eq. (7.21) is Hermitian.

Thus the functions $u(x)$ and $v(x)$ [solutions of Eq. (7.22)] must satisfy specified boundary conditions at the boundaries. These homogeneous boundary conditions can be of the types [$y(x)$ represents $u(x)$ or $v(x)$]

1. *Dirichlet boundary condition*:

$$y(a) = y(b) = 0.$$

2. *Neumann boundary condition*:

$$y'(a) = y'(b) = 0.$$

3. *Mixed boundary condition*:

$$y(a) + \alpha y'(a) = y(b) + \beta y'(b) = 0,$$

where α and β are specified non-vanishing constants. For the first two, Eqs. (7.24) and (7.25) are trivially satisfied. It can easily be seen that for the third type also Eq. (7.23) is satisfied. Hence for these boundary conditions, the operator $\hat{\mathcal{L}}$ in SL form is Hermitian.

Connection between physics and mathematics

We notice that conditions. (7.24) and (7.25) are also satisfied (and $\hat{\mathcal{L}}$ becomes Hermitian) if

$$p(a) = p(b) = 0, \tag{7.26}$$

i.e. if the end points of the interval become singular points, as seen by substituting Eq. (7.21) in Eq. (7.22). Since the physically acceptable solution must be well behaved *in the entire interval*, the physically allowed interval must be the region enclosed within the singular points, namely $a < x < b$. We call this interval the *mathematical interval*, while the interval defined by the potential as the *physical interval*. The condition (7.26) is frequently[1] satisfied by the actual *physical interval* in the differential equations encountered in quantum mechanics. We will see several examples later.

[1] This is true for the entire natural interval. For an artificially restricted physical interval, the end points need not be singular. An example is the infinite square well (Chap. 9, Sect. 9.1). In this case, the end points of the interval (chosen arbitrarily) are ordinary points and the physical condition demands that the wave function vanishes at the boundaries. Thus in this case, the operator $\hat{\mathcal{L}} = \frac{\mathrm{d}^2}{\mathrm{d}x^2}$ becomes Hermitian by Dirichlet boundary condition, *viz.* vanishing of the wave function at the boundaries.

As the boundary becomes a singular point, at least one solution diverges there. The diverging solution has to be discarded and the condition of regularity (finiteness of the physical solution) of the other solution determines the eigen value.

We will come back to the discussion of the relation of physics with mathematics after we discuss the properties satisfied by solutions of a differential eigen value equation.

Properties of (theorems on) eigen values and eigen functions of a differential Hermitian operator:
We list below important properties of eigen value equations satisfied by Hermitian differential operators. These (usually referred to as theorems) are the same as those for abstract Hermitian operators in a vector space or Hermitian matrices (see Chap. 2, Sects. 2.3.1 and 2.4). Proofs of the first two properties are provided below (see also Arfken 1966; Chattopadhyay 2006).

1. All eigen values are real.
2. Eigen functions belonging to different eigen values are orthogonal. Degenerate eigen functions (which belong to the *same eigen value*) are not automatically orthogonal, but can be orthogonalized.
3. The set of *all eigen functions* constitute a complete set.

Proof of the first two properties:
Consider the differential eigen value equation

$$\hat{\mathcal{L}}(x)u_{\lambda}(x) + \lambda w(x)u_{\lambda}(x) = 0, \qquad \text{[eq. (7.22)]}$$

Rewrite Eq. (7.22) for the eigen value λ'

$$\hat{\mathcal{L}}(x)u_{\lambda'}(x) + \lambda' w(x)u_{\lambda'}(x) = 0.$$

Taking complex conjugate we get (note that the differential operator is real since $p(x)$ and $q(x)$ of Eq. (7.21) are real functions. Also $w(x)$ is a real function)

$$\hat{\mathcal{L}}(x)u^*_{\lambda'}(x) + \lambda'^* w(x)u^*_{\lambda'}(x) = 0.$$

Pre-multiply Eq. (7.22) by $u^*_{\lambda'}(x)$ and the last equation by $u_{\lambda}(x)$, integrate both from $x = a$ to $x = b$ and subtract to get

$$\int_a^b u^*_{\lambda'}(x)\hat{\mathcal{L}}(x)u_{\lambda}(x)\mathrm{d}x - \int_a^b u_{\lambda}(x)\hat{\mathcal{L}}(x)u^*_{\lambda'}(x)\mathrm{d}x = (\lambda'^* - \lambda)\int_a^b u^*_{\lambda'}(x)u_{\lambda}(x)w(x)\mathrm{d}x.$$

Left side of the above equation vanishes by Eq. (7.23), since $\hat{\mathcal{L}}(x)$ is Hermitian. Hence

$$(\lambda'^{*} - \lambda) \int_a^b u^*_{\lambda'}(x) u_\lambda(x) w(x) \mathrm{d}x = 0. \tag{7.27}$$

Now setting $\lambda' = \lambda$, we get

$$(\lambda^* - \lambda) \int_a^b \left|u_\lambda(x)\right|^2 w(x) \mathrm{d}x = 0.$$

Since $w(x)$ is a non-vanishing positive function and $|u_\lambda(x)|^2$ is a positive function, the integral in the last equation can vanish *only if* $u_\lambda(x)$ vanishes identically (which is a trivial solution). Thus for non-trivial solutions, we must have

$$\lambda^* = \lambda,$$

i.e. all eigen values must be real, proving the first property. Hence Eq. (7.27) becomes

$$(\lambda' - \lambda) \int_a^b u^*_{\lambda'}(x) u_\lambda(x) w(x) \mathrm{d}x = 0.$$

For $\lambda' \neq \lambda$, we have

$$\int_a^b u^*_{\lambda'}(x) u_\lambda(x) w(x) \mathrm{d}x = 0 \quad \text{for } \lambda' \neq \lambda. \tag{7.28}$$

Thus eigen functions belonging to *different* eigen values are orthogonal.

However, it is possible that an eigen value λ appears more than once (called a degenerate eigen value), so that there are more than one eigen functions corresponding to the *same* eigen value. In this case we need other quantum numbers, say n, to distinguish the degenerate eigen functions $\{u_{\lambda,n}(x),$ with different $n\}$ belonging to a particular degenerate eigen value λ (see Chap. 8, Sect. 8.3). For such a situation, $u_{\lambda,n}(x)$ and $u_{\lambda,n'}(x)$ with $n \neq n'$ are not automatically orthogonal, but they are linearly independent (being *different* solutions of the *same* equation). Hence they can be orthogonalized (more generally orthonormalized) by a suitable procedure like the Gram–Schmidt method (see Chap. 2, Sect. 2.2.3).

Hence the set of *all eigen functions* of a Hermitian differential operator constitutes an orthogonal set. Furthermore since each one can be normalized, the set can be *orthonormal*.

Discussion of the third property (completeness): The third property asserts that any well behaved (i.e. continuous or at worst piecewise continuous and square-integrable) function $f(x)$ in $a \leq x \leq b$ and satisfying the same boundary conditions as $u_\lambda(x)$,

can be represented as a linear combination of the full set of eigen functions. Suppose the set is an infinite set, and we denote the infinite sequence of eigen values by $\{\lambda_m, m = 1, 2, \cdots\}$. Then $f(x)$ can be approximated by the series

$$f(x) = \sum_{m=1,2,\cdots}^{N} a_m u_{\lambda_m}(x),$$

(where a_m are constant coefficients) to any desired degree of accuracy when N is large enough. This means that the difference

$$\left| f(x) - \sum_{m=1,2,\cdots}^{N} a_m u_{\lambda_m}(x) \right|$$

can be made smaller than a pre-set small number ϵ, by increasing N.

If the set $\{u_{\lambda_m}(x)\}$ is orthonormal, the coefficient a_m is given by

$$a_m = \int_a^b f(x) u^*_{\lambda_m}(x) w(x) \mathrm{d}x.$$

If the set of eigen functions is *finite*, containing M eigen functions, then the sum is over all the M eigen functions, which span an M-dimensional vector space. For further discussion of completeness, see Chattopadhyay (2006).

Justification of the condition for a differential operator to be Hermitian as defined by Eq. (7.23):
Unlike a matrix operator, a differential operator in self-adjoint form (SL form) is not automatically Hermitian. This is because a differential equation does not produce a unique solution without a specified boundary condition. Hermitian adjoint of an abstract operator in a vector space is defined by Eq. (2.10). Replacing ψ, ϕ and $\hat{A}$ by v, u and $\hat{\mathcal{L}}$ respectively in complex conjugate the first equality of Eq. (2.10), and using the definition of an inner product, we have the condition for $\hat{\mathcal{L}}$ to be Hermitian [we leave out the argument (x) for brevity]

$$\begin{aligned}
\int_a^b v^* \hat{\mathcal{L}} u \mathrm{d}x &= (v, \hat{\mathcal{L}} u) = (u, \hat{\mathcal{L}}^\dagger v)^* \\
&= (u, \hat{\mathcal{L}} v)^* \quad \text{if } \hat{\mathcal{L}}^\dagger = \hat{\mathcal{L}} \text{ i.e. } \hat{\mathcal{L}} \text{ is Hermitian} \\
&= \left(\int_a^b u^* \hat{\mathcal{L}} v \mathrm{d}x \right)^* = \int_a^b u \hat{\mathcal{L}} v^* \mathrm{d}x \quad \text{(since } \hat{\mathcal{L}} \text{ is a real operator).}
\end{aligned}$$

Thus the differential operator $\hat{\mathcal{L}}$ is Hermitian if Eq. (7.23) is satisfied.

7.1.5 Connection Between Mathematics and Physics

In an n-dimensional system we have n variables. Then in a coordinate system which reflects the symmetry of the system, we can identify a radial variable (ρ), and $(n-1)$ remaining variables, collectively denoted by Ω. For example, for a two-dimensional system in polar coordinates, the variables are (r, θ) (with $x = r\cos\theta$ and $y = r\sin\theta$) and we identify $\rho \equiv r$, $\Omega \equiv \theta$. For a three-dimensional system in spherical polar coordinates, the variables are (r, θ, ϕ) and we identify $\rho \equiv r$, $\Omega \equiv (\theta, \phi)$.

For such a multi-dimensional system, the Schrödinger equation becomes a second order partial differential eigen value equation in (ρ, Ω). If the coordinate system corresponds to the symmetry of the physical system, the Schrödinger equation separates into single variable second order differential equations. One of them (usually the one in the radial variable ρ) involves energy in its eigen value term (see Sect. 9.3). This equation resembles an *effective one-dimensional Schrödinger equation*[2]:

$$\hat{h}(\rho)U_{\lambda}(\rho) = \lambda U_{\lambda}(\rho), \tag{7.29}$$

where $\hat{h}(\rho)$ is the effective one-dimensional Hamiltonian in the physical variable ρ, after separation of other variables (Ω) from the multi-dimensional (ρ, Ω) space. The effective Hamiltonian $\hat{h}(\rho)$ may contain terms arising from eigen values of remaining $(n-1)$ equations.

Corresponding eigen functions $\{U_{\lambda}(\rho)\}$ satisfy a physical orthogonality relation

$$\int_a^b U^*_{\lambda'}(\rho)U_{\lambda}(\rho)\xi(\rho)d\rho = 0 \quad \text{for } \lambda' \neq \lambda, \tag{7.30}$$

where the phase space factor $\xi(\rho)$ appears in the one-dimensional integral over ρ, arising from effects of other variables in the multi-dimensional volume element. For example, the volume element in two dimension is $d\tau = rdrd\theta \equiv \rho d\rho d\Omega$ and we identify $\xi(\rho) = \rho$. Similarly the volume element in three dimension is $d\tau = r^2 dr\sin\theta d\theta d\phi \equiv \rho^2 d\rho d\Omega$, and hence $\xi(\rho) = \rho^2$. The interval $a \leq \rho \leq b$ is obtained from the nature of the potential in $\hat{h}(\rho)$. This is the *physical interval*.

Eq. (7.29) is usually solved by transforming the variable ρ to a new variable $x = f(\rho)$ with a specific function $f(\rho)$. The interval in x is $a_x \leq x \leq b_x$, where $a_x = f(a)$ and $b_x = f(b)$. The interval $[a, b]$ and equivalently $[a_x, b_x]$ are physical

[2] However note that it is not the correct one-dimensional Schrödinger equation, for example in x-variable having $\frac{\mathrm{d}^2}{\mathrm{d}x^2}$. In three-dimensional spherical polar coordinates (r, θ, ϕ), the second derivative term for the radial wave function $R(r)$ in r-variable becomes $\frac{1}{r^2}\frac{\mathrm{d}}{dr}\left(r^2\frac{\mathrm{d}}{dr}\right)$. The latter can be reduced to $\frac{\mathrm{d}^2}{dr^2}$ for $\tilde{R}(r) = rR(r)$ [see Eq. (10.30)], which resembles a one-dimensional Schrödinger equation. But this is not really a one-dimensional Schrödinger equation as an extra centrifugal potential appears and the interval is $0 \leq r < \infty$ instead of $-\infty < r < \infty$ [see discussion following Eq. (10.30)].

intervals. Next, we investigate the asymptotic nature of the transformed equation in x, and obtain an asymptotic solution $g(x)$. We substitute

$$U_\lambda(\rho) = N_\lambda u_\lambda(x) g(x) \quad (\text{where } x = f(\rho), \text{ and } N_\lambda \text{ is a constant}), \tag{7.31}$$

in the differential eigen value equation in x, obtained from Eq. (7.29). Finally the resulting differential equation for $u_\lambda(x)$ is put in SL form to get

$$\hat{\mathcal{L}}(x) u_\lambda(x) = \lambda w(x) u_\lambda(x), \tag{7.32}$$

where $w(x)$ is the resulting weight function and λ is the eigen value. In many cases, the operator $\hat{\mathcal{L}}(x)$ can be identified as that of a standard mathematical function. The interval $[a', b']$ which makes $\hat{\mathcal{L}}(x)$ Hermitian according to Eq. (7.26), agrees with $[a_x, b_x]$, as we have already mentioned. This is the first connection between mathematics and physics.

Next the hermiticity of $\hat{\mathcal{L}}(x)$ in Eq. (7.32) guarantees a *mathematical orthogonality relation* [see Eq. (7.28)] in the interval $a' \le x \le b'$

$$\int_{a'}^{b'} u^*_{\lambda'}(x) u_\lambda(x) w(x) \mathrm{d}x = 0, \qquad \lambda' \neq \lambda \tag{7.33}$$

On the other hand, substituting Eq. (7.31) in the physical orthogonality relation (7.30), we have

$$\int_{a_x}^{b_x} u^*_{\lambda'}(x) u_\lambda(x) |g(x)|^2 \tilde{\xi}(x) \mathrm{d}x = 0, \qquad \lambda' \neq \lambda \tag{7.34}$$

where $\tilde{\xi}(x)$ comes from the transformation of the variable from ρ to x, including the function $\xi(\rho)$. In general it is found that besides $a' = a_x$ and $b' = b_x$, we also have $w(x) = |g(x)|^2 \tilde{\xi}(x)$. Thus the mathematical orthogonality relation agrees with the physical orthogonality relation. This is another connection between mathematics and physics. We present a few examples below. For further elaboration and examples, see Appendix A.

Ex. 1 Legendre differential equation

$$(1 - x^2) u''_\lambda(x) - 2x u'_\lambda(x) + \lambda u_\lambda(x) = 0$$

and its associated form (associated Legendre differential equation)

$$(1 - x^2) u''_\lambda(x) - 2x u'_\lambda(x) + \left(\lambda - \frac{m^2}{1 - x^2}\right) u_\lambda(x) = 0,$$

where m^2 is a real positive constant. In both these equations the eigen value is λ. These are the eigen value equations for the square of orbital angular momentum operator $(\hat{L}^2)$ in spherical polar coordinates (with $x = \cos\theta$). The first one is for axially symmetric case (see Chap. 10). Both the equations have regular singularities at $x = \pm 1$ and $x = \infty$. These are already in SL form and comparing with Eqs. (7.22) and (7.21), we have $p(x) = (1 - x^2)$ and $w(x) = 1$, while $q(x) = 0$ for the first and $q(x) = -\frac{m^2}{1-x^2}$ for the second equation. Hence $p(x) = 0$ gives $x = \pm 1$. Thus the corresponding differential operator is Hermitian in $[-1, 1]$. The natural (physical) interval for the polar angle θ is $[0, \pi]$, corresponding to $[-1, 1]$ for x, which coincides with the mathematical interval. Requirement for one of the solutions to be finite at the boundaries demands that its series solution must terminate (Chap. 10). This condition makes the eigen value quantized (for both equations)

$$\lambda = l(l+1) \quad \text{with} \quad l = 0, 1, 2, \cdots$$

The corresponding eigen functions $u_\lambda(x) = P_l(x)$ and $u^m_\lambda(x) = P^m_l(x)$ are called the Legendre polynomial of order l and associated Legendre function (not necessarily a polynomial) respectively.

Mathematical orthogonality relation for the Legendre polynomials

$$\int_{-1}^{1} P_{l'}(x) P_l(x) \mathrm{d}x = 0 \qquad \text{for } l' \neq l,$$

and for associated Legendre functions

$$\int_{-1}^{1} P^m_{l'}(x) P^m_l(x) \mathrm{d}x = 0 \qquad \text{for } l' \neq l,$$

are the same as physical (quantum mechanical) orthogonality relations (see Chap 10).

Ex. 2. Hermite equation:

$$u''_\lambda(x) - 2x\, u'_\lambda(x) + 2\lambda\, u_\lambda(x) = 0,$$

where λ is a constant (2λ is the eigen value). This equation appears in the Schrödinger equation for the one-dimensional harmonic oscillator (Sect. 9.4, Chap. 9). It has irregular singularities at $x = \pm\infty$ and is not in SL form. It can be put in SL form by multiplying it with e^{-x^2}

$$\mathrm{e}^{-x^2} u''_\lambda(x) - 2x\, \mathrm{e}^{-x^2} u'_\lambda(x) + 2\lambda\, \mathrm{e}^{-x^2} u_\lambda(x) = 0.$$

Comparing with Eqs. (7.22) and (7.21), we identify $p(x) = \mathrm{e}^{-x^2}$, $q(x) = 0$, $w(x) = \mathrm{e}^{-x^2}$ and eigen value 2λ. Here $p(x) = 0$ gives $x = \pm\infty$. Hence the oper-

ator is Hermitian in the interval $-\infty < x < \infty$. This coincides with the physical interval for a one-dimensional harmonic oscillator. For an infinite series solution about $x = 0$ the complete wave function diverges at the boundaries (in agreement with the singularities at $x = \pm\infty$), unless the series is terminated to a polynomial by the condition

$$\lambda = n, \quad \text{with} \quad n = 0, 1, 2, \cdots$$

Thus the eigen value is quantized. Corresponding eigen function of the Hermite equation is called Hermite polynomial $[H_n(x)]$ and satisfies the mathematical orthogonality relation

$$\int_{-\infty}^{\infty} H_{n'}(x) H_n(x) \mathrm{e}^{-x^2} \mathrm{d}x = 0 \qquad \text{for } n' \neq n.$$

For the physical wave function of the one-dimensional harmonic oscillator, we have $g(x) = \mathrm{e}^{-\frac{1}{2}x^2}$ and $\tilde{\xi}(x) = 1$ (Sect. 9.4, Chap. 9). Hence $w(x) = |g(x)|^2 \tilde{\xi}(x)$, which shows that the physical and mathematical orthogonality relations agree.

Ex. 3. Laguerre equation:

$$x\, u_\lambda''(x) + (1 - x)\, u_\lambda'(x) + \lambda\, u_\lambda(x) = 0$$

and its associated form (associated Laguerre differential equation)

$$x\, u_\lambda''(x) + (P + 1 - x)\, u_\lambda'(x) + \lambda\, u_\lambda(x) = 0 \quad (P \text{ is a real positive constant})$$

[Symbol λ for the eigen value has been retained to be consistent with its previous use in this chapter, although it is different from the standard use in Eq. (7.52)]. These equations appear in the quantum mechanical treatment of the hydrogen atom (Chap. 11) and 3-D harmonic oscillator (Sect. 12.4, Chap. 12). Laguerre equation is a special case $(P = 0)$ of the more general associated Laguerre equation. Both these equations have one regular singularity at $x = 0$ and an irregular singularity at $x = \infty$. They are not in SL form and can be put in that form by multiplying with e^{-x} and $x^P \mathrm{e}^{-x}$ respectively:

$$x\, \mathrm{e}^{-x} u_\lambda''(x) + (1 - x)\, \mathrm{e}^{-x} u_\lambda'(x) + \lambda\, \mathrm{e}^{-x} u_\lambda(x) = 0,$$

$$\text{and} \quad x^{P+1} \mathrm{e}^{-x} u_\lambda''(x) + (P + 1 - x) x^P \mathrm{e}^{-x} u_\lambda'(x) + \lambda\, x^P \mathrm{e}^{-x} u_\lambda(x) = 0.$$

From these, we can identify

$p(x) = x\mathrm{e}^{-x}$, $q(x) = 0$, $w(x) = \mathrm{e}^{-x}$ for Laguerre equation

and

$p(x) = x^{P+1}\mathrm{e}^{-x}$, $q(x) = 0$, $w(x) = x^P \mathrm{e}^{-x}$ for associated Laguerre equation.

Eigen value is λ for both the equations. For both, $p(x)$ vanishes for $x = 0$ and ∞. Thus corresponding differential operators are Hermitian in the interval $0 \leq x < \infty$, which again coincides with the natural physical interval for both H-atom and 3-D harmonic oscillator.

For both equations, if a series solution about $x = 0$ is allowed to be an infinite series, it diverges as $x \to \infty$ (in agreement with the fact that $x = \infty$ is an irregular singularity). This series can be terminated into a polynomial [and hence the full wave function (which contains a factor $e^{-\frac{x}{2}}$) remains finite even as $x \to \infty$] by choosing

$$\lambda = n, \quad n = 1, 2, 3, \cdots .$$

This results in energy quantization. Corresponding eigen function of the associated Laguerre equation $[L_\lambda^P(x)]$ satisfies the mathematical orthogonality relation with $w(x) = x^P e^{-x}$ (also valid for Laguerre polynomials with $P = 0$)

$$\int_0^\infty L_{\lambda'}^P(x) L_\lambda^P(x) x^P e^{-x} dx = 0 \qquad \text{for } \lambda' \neq \lambda.$$

This agrees with the physical orthogonality relation for the 3-D harmonic oscillator, but not for the H-atom (See Appendix A for details).

7.2 Some Standard Differential Equations

In this section, we list a number of commonly encountered standard differential equations, together with their standard solutions and some properties (see Arfken 1966; Chattopadhyay 2006). Some of these have already been discussed in connection with hermiticity. These equations arise often in physical problems, including those in quantum mechanics. Hence this list will be quite useful in our discussions. In the following x is the variable of the differential equation and $y(x)$ is its solution. Real constants appearing in the equation are specified.

1. **Simple harmonic equation**

$$y''(x) + \alpha^2 y(x) = 0 \quad (\alpha^2 \text{ is a real positive constant}). \tag{7.35}$$

For this equation, there are no singularities for finite x, but irregular singularities exist at $x = \pm\infty$. Hence the mathematical interval is $-\infty < x < \infty$, unless restricted by the problem (e.g. a particle in a square well with infinitely rigid walls at $x = \pm a$, see Chap. 9). The equation is already in SL form, and we identify: $p(x) = 1$, $q(x) = 0$, $\lambda = \alpha^2$ and $w(x) = 1$. Since in this case $p(x) = 1$, Eq. (7.26) cannot be used for hermiticity. Instead Dirichlet or Neumann boundary conditions are to be used, which give rise to discrete eigen values

for α. The standard solutions $\sin\alpha x$ and $\cos\alpha x$ (or equivalently $e^{i\alpha x}$ and $e^{-i\alpha x}$) are oscillating functions, well behaved in the interval $-\infty < x < \infty$. Hence both can appear in the general solution. Boundary conditions (Dirichlet for $\sin\alpha x$ or Neumann for $\cos\alpha x$), which make the operator $\frac{d^2}{dx^2}$ Hermitian, provide the mathematical interval as $0 \le x \le 2\pi$. This results in discrete eigen values: $\alpha = n$, where n is an integer. For $n = 0$ the solution $\sin nx$ becomes trivial. Physical restrictive conditions (Dirichlet boundary conditions), as in the case of an infinite square well in one dimension (see Sect. 9.1, Chap. 9), give energy eigen values in terms of α..

The solutions $\sin nx$ and $\cos nx$ satisfy well known orthonormality relations (Arfken 1966), in the interval $[0, 2\pi]$ for integral n, n':

$$\begin{aligned}
\int_0^{2\pi} \sin n'x \ \sin nx \ dx &= \pi\,\delta_{n',n}, \quad (\text{for } n' \neq 0) \\
&= 0, \quad (\text{for } n' = 0 \text{ or } n = 0). \\
\int_0^{2\pi} \cos n'x \ \cos nx \ dx &= \pi\,\delta_{n',n}, \quad (\text{for } n' \neq 0) \\
&= 2\pi, \quad (\text{for } n' = n = 0). \\
\int_0^{2\pi} \sin n'x \ \cos nx \ dx &= 0, \quad (\text{for all integral } n, n').
\end{aligned} \tag{7.36}$$

If α^2 is negative (such that $\alpha^2 = -\beta^2$, with β real positive), the linearly independent solutions are $\sinh\beta x$ and $\cosh\beta x$, or equivalently $e^{\beta x}$ and $e^{-\beta x}$. The solution $e^{\beta x}$ vanishes at $x = -\infty$ and diverges at $x = \infty$, while $e^{-\beta x}$ diverges at $x = -\infty$ and vanishes at $x = \infty$. Since these are not orthogonal, they cannot appear as complete solutions, but can appear as a solution in a part of the interval, as in the case of a finite square well (see Sect. 9.2, Chap. 9).

2. **Hermite differential equation**

$$y''(x) - 2xy'(x) + 2ay(x) = 0 \quad (a \text{ is a real constant}). \tag{7.37}$$

This equation has no singularities for finite x and irregular singularities at $x = \pm\infty$. It appears in the quantum mechanical treatment of one-dimensional linear harmonic oscillator (see Sect. 9.4, Chap. 9), for which the physical interval is $-\infty < x < \infty$.

Eq. (7.37) is not in SL form and multiplication through by e^{-x^2} puts it in that form

$$e^{-x^2} y''(x) - 2x\,e^{-x^2} y'(x) + 2a\,e^{-x^2} y(x) = 0,$$

with $p(x) = \mathrm{e}^{-x^2}$, $q(x) = 0$, $\lambda = 2a$ and $w(x) = \mathrm{e}^{-x^2}$. Now $p(x)$ vanishes at $x = \pm\infty$. Hence the mathematical interval is $-\infty < x < \infty$. Standard linearly independent solutions are called Hermite functions. Series solutions for both these functions diverge for $|x| \to \infty$, if allowed to be infinite series. One of these terminates and becomes a polynomial (called Hermite polynomial $H_n(x)$) for the choice $a = n$ (a non-negative integer), while the other series does not terminate for this choice of a, but diverges as $|x| \to \infty$, and should be discarded. The polynomial, called Hermite polynomial, is analytic in the entire interval and is the physically acceptable solution, while the diverging solution is rejected.
If for a specific case, the physical interval is restricted to a finite one, both solutions will contribute, with appropriate boundary conditions.
A two-variable generating function $g(x, t)$, which defines the Hermite polynomials is given by (Arfken 1966)

$$g(x,t) = \mathrm{e}^{-t^2+2xt} = \sum_{n=0}^{\infty} H_n(x)\,\frac{t^n}{n!}. \tag{7.38}$$

Expanding the exponential and comparing coefficients of equal powers of t on both sides, one can obtain a power series expansion for $H_n(x)$. On the other hand differentiating both sides w.r.t. t and x we get respectively the following recurrence relations

$$\begin{aligned} H_{n+1}(x) &= 2x\,H_n(x) - 2n\,H_{n-1}(x), \\ H_n'(x) &= 2n\,H_{n-1}(x). \end{aligned} \tag{7.39}$$

Properties of $H_n(x)$ can be obtained from these equations.
The orthonormality relation of the standard Hermite polynomials is (Chattopadhyay 2006) [note that the weight function $w(x) = \mathrm{e}^{-x^2}$ and mathematical interval $-\infty < x < \infty$ guarantee the orthogonality relation]

$$\int_{-\infty}^{\infty} H_{n'}(x)\,H_n(x)\mathrm{e}^{-x^2}\mathrm{d}x = \sqrt{\pi}\,2^n\,n!\,\delta_{n,n'}. \tag{7.40}$$

3. **Legendre differential equation**

$$(1-x^2)y''(x) - 2xy'(x) + l(l+1)y(x) = 0 \quad (l \text{ is a real constant}). \tag{7.41}$$

This equation appears as the quantum mechanical eigen value equation for $\hat{L}^2$ (orbital angular momentum squared) operator in spherical polar coordinates, with axial symmetry (Chap. 10). It has regular singularities at $x = \pm 1$ and $x = \infty$ and is already in SL form, with $p(x) = (1 - x^2)$, $q(x) = 0$, $\lambda = l(l + 1)$ and $w(x) = 1$. The mathematical interval is $-1 \le x \le 1$, as $p(x) = 0$ gives $x = \pm 1$.

This coincides with the unrestricted physical interval of polar angle θ for the $\hat{L}^2$-equation, which is $0 \leq \theta \leq \pi$ corresponding to $-1 \leq x \leq 1$, where $x = \cos\theta$. Standard linearly independent solutions are: Legendre functions of first and second kinds. Both these diverge at $x = \pm 1$. One solution can be converted to a polynomial [called Legendre polynomial, $P_l(x)$] by choosing l to be a non-negative integer, which is the acceptable solution. For this choice, the other solution diverges at $|x| = 1$ and is called Legendre function of second kind, $Q_l(x)$. Thus $Q_l(\theta)$ diverges along the z-axis. Usually only $P_l(x)$ are needed for a physical solution. However, if a specific case restricts the physical interval to exclude the z-axis, such that $-x_0 \leq x \leq x_0$ with $|x_0| < 1$, both $P_l(x)$ and $Q_l(x)$ can contribute.
The generating function for the Legendre polynomials is (Arfken 1966)

$$g(x,t) = \left(1 - 2xt + t^2\right)^{-\frac{1}{2}} = \sum_{l=0}^{\infty} P_l(x)\, t^l. \tag{7.42}$$

From this equation other properties (like series expansion and recurrence relations) of Legendre polynomials can be derived. The recurrence relations for integral $l \geq 1$ are (Arfken 1966)

$$\begin{aligned} &(2l+1)x\,P_l(x) = (l+1)P_{l+1}(x) + l\,P_{l-1}(x), \\ &P'_{l+1}(x) - P'_{l-1}(x) = (2l+1)P_l(x). \end{aligned} \tag{7.43}$$

Since $w(x) = 1$ and mathematical interval is $-1 \leq x \leq 1$, the orthonormality relation satisfied by the standard Legendre polynomials is (Chattopadhyay 2006)

$$\int_{-1}^{1} P_{l'}(x)\, P_l(x)\mathrm{d}x = \frac{2}{2l+1}\delta_{l',l}. \tag{7.44}$$

4. **Associated Legendre differential equation**

$$(1-x^2)y''(x) - 2xy'(x) + \left[l(l+1) - \frac{m^2}{1-x^2}\right]y(x) = 0$$
$$(l, m \text{ are real constants}). \tag{7.45}$$

Its singularities, mathematical interval, eigen value and weight function are the same as those of the Legendre equation, while $q(x) = -\frac{m^2}{1-x^2}$ such that m is a constant of the equation. This equation appears in the general treatment of square of orbital angular momentum operator in spherical polar coordinates (Chap. 10). Standard solutions are the associated Legendre functions of first and second kinds. Although the quantity m is restricted to all integers from the uniqueness condition of the ϕ-solution, acceptability of the full solution of Eq. (7.45) requires m to be restricted to integers $-l \leq m \leq l$, with l non-negative

integer. When these conditions are satisfied, one of the two linearly independent solutions becomes acceptable (although not a polynomial for odd integral m) and is called associated Legendre function, $P_l^m(x)$ (see Chap. 10, Sect. 10.2). The other solution $Q_l^m(x)$, (called associated Legendre function of second kind) diverges at $|x| = 1$ and does not appear in the physical solution, unless the z-axis is excluded from the physical domain. Note that Eq. (7.45) remains unchanged when m is replaced by $-m$. Hence $P_l^{-m}(x)$ is proportional to $P_l^m(x)$. For $m = 0$, we get Legendre polynomial: $P_l^0(x) = P_l(x)$.
The associated Legendre equation can be obtained from the Legendre equation by differentiating it m-times (using Leibnitz formula for multiple differentiation of a product of two functions) (Arfken 1966). Hence associated Legendre functions can be obtained from Legendre polynomials

$$P_l^m(x) = (1 - x^2)^{\frac{m}{2}} \frac{\mathrm{d}^m}{\mathrm{d}x^m} P_l(x). \tag{7.46}$$

A generating function can be defined in this case, but it is not very useful. This function can be used to get recurrence relations. However, since there are two indices (l and m) there will be several recurrence relations. Recurrence relations can also be obtained from those of Legendre polynomial by using Eq. (7.46). Since these are not much used, we do not present them here, but can be found in Arfken (1966).
The orthonormality relation satisfied by the standard $P_l^m(x)$ is (Arfken 1966) [note again that the weight function $w(x) = 1$ and the mathematical interval is $-1 \le x \le 1$]

$$\int_{-1}^{1} P_{l'}^m(x)\, P_l^m(x)\, \mathrm{d}x = \frac{2}{2l+1} \frac{(l+m)!}{(l-m)!} \delta_{l',l}. \tag{7.47}$$

This reduces to Eq. (7.44) for $m = 0$.

5. **Laguerre differential equation**

$$xy''(x) + (1 - x)y'(x) + \lambda y(x) = 0 \quad (\lambda \text{ is a real constant}). \tag{7.48}$$

This is a special case of the associated Laguerre differential equation (see below) and has a regular singularity at $x = 0$ and an irregular singularity at $x = \infty$. Multiplication through by e^{-x} puts this equation in SL form. We identify $p(x) = x\mathrm{e}^{-x}, q(x) = 0, \lambda$ is the eigen value and $w(x) = \mathrm{e}^{-x}$. Hence solutions of $p(x) = 0$ give the mathematical interval as $0 \le x < \infty$. This equation appears for the $l = 0$ states of H-atom (Chap. 11), for which the physical interval coincides with the mathematical interval. Out of the two linearly independent solutions (called Laguerre functions of first and second kinds), one becomes a polynomial (hence analytic in the entire interval) when λ is a non-negative integer, say n. This

polynomial is called Laguerre polynomial $L_n(x)$. The condition $\lambda = n$ leads to quantization.
The generating function is (Arfken 1966)

$$g(x,z) = \frac{\mathrm{e}^{-\frac{xz}{1-z}}}{1-z} = \sum_{n=0}^{\infty} L_n(x) z^n, \quad |z| < 1. \tag{7.49}$$

Differentiation w.r.t. x and z give the recurrence relations (Arfken 1966)

$$\begin{aligned} (n+1)L_{n+1}(x) &= (2n+1-x)L_n(x) - nL_{n-1}(x), \\ xL_n'(x) &= nL_n(x) - nL_{n-1}(x). \end{aligned} \tag{7.50}$$

The mathematical interval $0 \le x < \infty$ and the weight function $w(x) = \mathrm{e}^{-x}$ are consistent with the orthonormality relation for the standard Laguerre polynomials (Arfken 1966)

$$\int_0^{\infty} L_{n'}(x)\, L_n(x)\, \mathrm{e}^{-x}\, \mathrm{d}x = \delta_{n',n} \tag{7.51}$$

6. **Associated Laguerre differential equation**

$$\begin{aligned} xy''(x) + (P+1-x)y'(x) + (Q-P)y(x) = 0 \\ (P, Q \text{ are real constants}). \end{aligned} \tag{7.52}$$

With $P = 0$ (and renaming Q as λ) this equation reduces to Laguerre equation (7.48). The singularities are the same as those for the Laguerre equation. Multiplication through by $x^P \mathrm{e}^{-x}$ puts this equation in SL form, with the identification: $p(x) = x^{P+1}\mathrm{e}^{-x}$, $q(x) = 0$, $\lambda = Q - P$ and weight function $w(x) = x^P \mathrm{e}^{-x}$. Since $p(x) = 0$ at $x = 0$ and $x = \infty$ the mathematical interval is $0 \le x < \infty$. The standard solution is called the associated Laguerre function, $L_Q^P(x)$. A series solution about $x = 0$ diverges as $x \to \infty$, unless it is terminated to a polynomial of degree n by the condition $Q - P = n$, where n is a non-negative integer. This polynomial is called the associated Laguerre polynomial. The associated Laguerre equation results from the radial Schrödinger equation for H-type atoms with non-vanishing orbital angular momentum, l. The natural physical interval coincides with the mathematical interval. For the solution for an arbitrary $l \ge 0$ to be acceptable one has to choose $P = 2l + 1$ and $Q = n + l$ with n an integer ≥ 1, which leads to energy quantization (Chap. 11). We will see another use in Chap. 12 in connection with the three-dimensional harmonic oscillator in spherical polar coordinates.
As in the case of Legendre equations, the associated Laguerre equation can be obtained from the Laguerre equation by P-times differentiation of the Laguerre equation for $L_{n+P}(x)$, with the definition (note that this definition is not unique (Arfken 1966)

$$L_n^P(x) = (-1)^P \frac{\mathrm{d}^P}{\mathrm{d}x^P} L_{n+P}(x). \tag{7.53}$$

This shows that $L_n^0(x) = L_n(x)$. We can also obtain a generating function for the associated Laguerre polynomials from that for Laguerre polynomials

$$g_P(x,z) = \frac{\mathrm{e}^{-\frac{xz}{1-z}}}{(1-z)^{P+1}} = \sum_{n=0}^{\infty} L_n^P(x) z^n, \quad |z| < 1. \tag{7.54}$$

A number of recurrence relations can be derived from this generating function, which can be found in Arfken (1966).
The weight function $w(x) = x^P \mathrm{e}^{-x}$ and mathematical interval $0 \le x < \infty$ are consistent with the orthonormality relation (Arfken 1966)

$$\int_0^{\infty} L_{n'}^P(x)\, L_n^P(x)\, x^P \mathrm{e}^{-x} \mathrm{d}x = \frac{(n+P)!}{n!}\, \delta_{n',n}. \tag{7.55}$$

See Appendix A for comparison of mathematical and physical orthogonality relations.

7. **Bessel differential equation**

$$x^2 y''(x) + x y'(x) + (x^2 - \nu^2) y(x) = 0 \quad (\nu \text{ is a real constant}). \tag{7.56}$$

This equation appears in problems with cylindrical symmetry, for the radial motion in a plane perpendicular to the axis of symmetry (see Sect. 12.3, Chap. 12). A two-dimensional problem with circular symmetry is a special case, as this can be considered as a section of a right circular cylinder. As for an example, the classical vibration of a circular membrane results in this equation. Bessel equation has a regular singularity at $x = 0$ and an irregular singularity at $x = \infty$. Hence the unrestricted mathematical interval is $0 \le x < \infty$, which again coincides with the unrestricted physical domain of the problem. For restricted physical domain, this equation can be converted into an eigen value equation for a fixed ν (see below and Sect. 12.3, Chap. 12). The linearly independent solutions are the standard Bessel function, $J_\nu(x)$ and the Neumann function, $N_\nu(x)$ (both of order ν). Neither of these is a polynomial. Only $J_\nu(x)$ is finite at $x = 0$, while $N_\nu(x)$ diverges at the origin. Hence for physical problems that include the origin (actually the axis of the right circular cylinder), only $J_\nu(x)$ will contribute. However if the axis is excluded (e.g., a quantum mechanical particle confined in the space between two concentric right circular cylinders of radii R_1 and R_2 and height H, having rigid walls on all sides), we have to include both solutions.
Note that the constant ν need not be an integer, unless required by another physical requirement. It can be positive or negative. In fact $J_\nu(x)$ and $J_{-\nu}(x)$ are linearly independent solutions, *unless* ν *is an integer*. The Neumann function

$N_\nu(x)$ is a linear combination of $J_\nu(x)$ and $J_{-\nu}(x)$ defined as (Arfken 1966)

$$N_\nu(x) = \frac{\cos \nu\pi\, J_\nu(x) - J_{-\nu}(x)}{\sin \nu\pi}. \tag{7.57}$$

Two other linear combinations of $J_\nu(x)$ and $N_\nu(x)$ called Hankel function of first ($H_\nu^{(1)}(x)$) and second ($H_\nu^{(2)}(x)$) kind are also used as the linearly independent solutions of Eq. (7.56) (Arfken 1966)

$$\begin{aligned} H_\nu^{(1)}(x) &= J_\nu(x) + i N_\nu(x), \\ H_\nu^{(2)}(x) &= J_\nu(x) - i N_\nu(x). \end{aligned} \tag{7.58}$$

These are useful for quantum mechanical traveling wave with cylindrical symmetry. These definitions are analogous to $\mathrm{e}^{\pm ix} = \cos x \pm i \sin x$.
Eq. (7.56) is apparently not an eigen value equation for a *fixed constant* ν. However replacing x by $\frac{\zeta_\nu}{a}x$ (where ζ_ν and a are constants, to be specified later) and dividing by x, we have for the solution $y(x)$ replaced by $J_\nu(\frac{\zeta_\nu}{a}x)$

$$x\frac{\mathrm{d}^2 J_\nu(\frac{\zeta_\nu}{a}x)}{\mathrm{d}x^2} + \frac{\mathrm{d}J_\nu(\frac{\zeta_\nu}{a}x)}{\mathrm{d}x} + x\Big(\frac{\zeta_\nu^2}{a^2} - \frac{\nu^2}{x^2}\Big) J_\nu\Big(\frac{\zeta_\nu}{a}x\Big) = 0. \tag{7.59}$$

Then we can identify the quantity $\frac{\zeta_\nu^2}{a^2}$ as the eigen value. This equation is in SL form with

$$p(x) = x, \quad q(x) = -\frac{\nu^2}{x}, \quad \lambda = \frac{\zeta_\nu^2}{a^2}, \quad \text{and} \quad w(x) = x.$$

We next have to make sure that the corresponding differential operator is Hermitian. Since $p(x) = 0$ has only one solution, *viz.* $x = 0$, we can choose one extremity of the interval at $x = 0$. Then choosing the other extremity as $x = a$, the condition for hermiticity demands (Dirichlet boundary condition)

$$J_\nu\Big(\frac{\zeta_\nu}{a}x\Big)\Big|_{x=a} = 0, \quad \text{i.e.} \quad J_\nu(\zeta_\nu) = 0. \tag{7.60}$$

Thus ζ_ν is a zero of the Bessel function $J_\nu(x)$. Let it be the n-th zero $\zeta_{\nu,n}$. Hence the eigen value of Eq. (7.59) is $\frac{\zeta_{\nu,n}^2}{a^2}$, the eigen value index being n. The complete mathematical orthonormality relation is (see Eq. (5.46) of Chattopadhyay 2006)

$$\int_0^a J_\nu\Big(\frac{\zeta_{\nu,n}}{a}x\Big) J_\nu\Big(\frac{\zeta_{\nu,m}}{a}x\Big)\, x\, \mathrm{d}x = \frac{a^2}{2}\big[J_{\nu+1}(\zeta_{\nu,n})\big]^2 \delta_{n,m}\,, \tag{7.61}$$

where the factor x in the integrand comes from the weight function.

The generating function for Bessel function $J_\nu(x)$ is (Arfken 1966)

$$g(x,t) = \mathrm{e}^{\frac{x}{2}(t-\frac{1}{t})} = \sum_{n=-\infty}^{\infty} J_n(x)t^n. \tag{7.62}$$

It can be used to get a series solution and recurrence relations for $J_n(x)$. The recurrence relations are

$$\begin{aligned} J_{n-1}(x) + J_{n+1}(x) &= \frac{2n}{x} J_n(x) \\ J_{n-1}(x) - J_{n+1}(x) &= 2\, J_n'(x). \end{aligned} \tag{7.63}$$

Neumann functions satisfy the same recurrence relations (Arfken 1966).

8. **Modified Bessel equation**

$$x^2y''(x) + xy'(x) - (x^2 + \nu^2)y(x) = 0 \quad (\nu \text{ is a real constant}). \tag{7.64}$$

This equation appears in problems with cylindrical symmetry, with a distributed sink term, *e.g.*, an absorption due to scattering of a neutron flux. The variable x is the perpendicular distance of a point from the axis of the cylinder.
Eq. (7.64) has the same singularities and unrestricted mathematical interval as those of the Bessel equation. It is similar to Bessel equation (7.56), except for the sign of x^2 in front of $y(x)$ and is obtained by replacing x by ix in Eq. (7.56). Thus the solution of Eq. (7.64), called the modified Bessel function of first kind (with an arbitrary, but convenient constant $i^{-\nu}$), is (Arfken 1966)

$$I_\nu(x) = i^{-\nu} J_\nu(ix). \tag{7.65}$$

As in the case of Bessel functions, $I_\nu(x)$ and $I_{-\nu}(x)$ are linearly independent solutions of the modified Bessel equation, *except when* ν *is an integer*, when these are identical. Similar to Neumann function, a second linearly independent solution of Eq. (7.64) is defined as (Arfken 1966)

$$K_\nu(x) = \frac{\pi}{2} \frac{I_{-\nu}(x) - I_\nu(x)}{\sin \nu\pi}, \tag{7.66}$$

(once again with an arbitrary but convenient constant $\frac{\pi}{2}$) and is called the modified Bessel function of second kind. Note that the definitions (7.65) and (7.66) are not unique. We present the definitions given in Arfken (1966).
$I_\nu(x)$ is finite at $x = 0$ (along the axis of the cylinder), but diverges as x increases (away from the axis). On the other hand, $K_\nu(x)$ decreases as x increases. but diverges at $x = 0$. Coefficients of a general solution [a linear combination of $I_\nu(x)$ and $K_\nu(x)$] are determined from the boundary conditions along the axis and away from it. Plots of integral order functions, *viz.* $I_n(x)$ and $K_n(x)$ can be

found in Arfken (1966). Note that both types of modified Bessel functions are monotonic functions of x.
Generating function for integral order $I_n(x)$ is given by (Chattopadhyay 2006)

$$g(x,t) = \mathrm{e}^{\frac{x}{2}\left(t+\frac{1}{t}\right)} = \sum_{n=-\infty}^{\infty} I_n(x)t^n. \tag{7.67}$$

Recurrence relations can be obtained from the generating function (or from those for $J_\nu(x)$) as (Arfken 1966)

$$\begin{aligned} I_{\nu-1}(x) - I_{\nu+1}(x) &= \frac{2\nu}{x} I_\nu(x), \\ I_{\nu-1}(x) + I_{\nu+1}(x) &= 2\, I'_\nu(x). \end{aligned} \tag{7.68}$$

Unlike $J_n(x)$, modified Bessel functions of first kind $[I_n(x)]$ do not satisfy any orthogonality relation, like the exponential functions $\mathrm{e}^{\pm bx}$, which are the solutions of the equation $y''(x) - b^2 y(x) = 0$. Note that $I_n(x)$ and $K_n(x)$ are not oscillating, but monotonic functions (see plots of $n = 0$ and $n = 1$ in Arfken 1966). We can see that $I_n(x)$ cannot satisfy an orthogonality relation, as follows. We can proceed exactly as in Eq. (7.59) and expect an orthogonality relation like Eq. (7.61). Hermiticity condition of the operator requires the upper limit of the mathematical interval, as well as the upper limit of the orthogonality integral to be a zero of $I_n(x)$ [as in the case for $J_n(x)$]. But the only zero of $I_n(x)$ is $x = 0$ for $n \neq 0$ ($I_0(x)$ has no zeros, see Arfken (1966), causing the interval to vanish, and the operator to be non-Hermitian. Indeed Eq. (7.64) does not correspond to a physical eigen value equation.

9. **Spherical Bessel equation**

$$x^2 y''(x) + 2x\, y'(x) + \left[x^2 - l(l+1)\right] y(x) = 0. \tag{7.69}$$

(l is an integral constant)

This equation appears in problems with spherical symmetry (or with conical symmetry, in which θ is restricted to $\theta \leq \theta_c$ by the condition of the problem). Here x is the distance from the origin (the r variable in spherical polar coordinates). The constant l represents the orbital angular momentum and is restricted to non-negative integers by the angular momentum quantization. An example of its application is the quantum mechanical problem of a particle in a spherical hole with a rigid wall (see Sect. 12.1, Chap. 12).
Equation (7.69) is already in SL form and has the same singularities as the Bessel equation. Hence the complete mathematical interval is $0 \leq x < \infty$, which agrees with the unrestricted physical interval (the latter can be restricted by the problem, see below). For a fixed value of l this equation apparently does not have an eigen value term. However we can have an eigen value as a scaling factor in a restricted interval (as in the case of Bessel function), as we will see below.

Equation (7.69) can be put in the form of Bessel equation by the substitution $y(x) = \frac{Z_l(x)}{\sqrt{x}}$, where $Z_l(x)$ is either $J_{l+\frac{1}{2}}(x)$ or $N_{l+\frac{1}{2}}(x)$. Because of their frequent occurrence, these functions are separately named as *spherical Bessel function* and *spherical Neumann function* respectively and are defined (with arbitrary but convenient constants) as (Arfken 1966)

$$\begin{aligned} \text{Spherical Bessel function :} \quad & j_l(x) = \sqrt{\frac{\pi}{2x}} J_{l+\frac{1}{2}}(x), \\ \text{Spherical Neumann function :} \quad & n_l(x) = \sqrt{\frac{\pi}{2x}} N_{l+\frac{1}{2}}(x). \end{aligned} \tag{7.70}$$

Thus the standard linearly independent solutions of Eq. (7.69) are $j_l(x)$ and $n_l(x)$, which are not polynomials. These have properties similar to those of Bessel and Neumann functions as can be inferred from Eq. (7.70). Like $N_l(x)$, spherical Neumann function $n_l(x)$ also diverges at $x = 0$, while $j_l(x)$ is finite at $x = 0$. Both these functions oscillate with gradually decreasing amplitude as x increases and each has an infinite number of discrete zeros. These functions can be expressed in terms of elementary trigonometric functions (see Arfken 1966). As in the case of Bessel functions, two other linear combinations of $j_l(x)$ and $n_l(x)$ are commonly used and are called spherical Hankel function of first and second kinds (Arfken 1966)

$$\begin{aligned} h_l^{(1)}(x) &= j_l(x) + i n_l(x), \\ h_l^{(2)}(x) &= j_l(x) - i n_l(x). \end{aligned} \tag{7.71}$$

These have recurrence relations and other properties which can be inferred from those of spherical Bessel functions and are useful in scattering problems.
Two very useful general recurrence relations for all the spherical Bessel functions can be expressed as (Arfken 1966)

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}\left[x^{l+1} f_l(x)\right] &= x^{l+1} f_{l-1}(x), \\ \frac{\mathrm{d}}{\mathrm{d}x}\left[x^{-l} f_l(x)\right] &= -x^{-l} f_{l+1}(x), \end{aligned} \tag{7.72}$$

where $f_l(x)$ represents any one of the spherical Bessel functions, *viz.* $j_l(x)$, $n_l(x)$, $h_l^{(1)}(x)$ or $h_l^{(2)}(x)$.
Eigen value and orthogonality relation: as in the case of the Bessel equation, the spherical Bessel equation also does not have an apparently clear eigen value term for a fixed l. Substituting

$$x = \frac{z_{l,m}}{a}\rho, \qquad \text{where} \quad z_{l,m} \quad \text{and} \quad a \quad \text{are constants,}$$

in Eq. (7.69) and writing $y(x) = j_l\left(\frac{z_{l,m}}{a}\rho\right)$ we get

$$\rho^2 \frac{\mathrm{d}^2}{d\rho^2} j_l\left(\frac{z_{l,m}}{a}\rho\right) + 2\rho \frac{\mathrm{d}}{d\rho} j_l\left(\frac{z_{l,m}}{a}\rho\right) + \left[\frac{(z_{l,m})^2}{a^2}\rho^2 - l(l+1)\right] j_l\left(\frac{z_{l,m}}{a}\rho\right) = 0. \tag{7.73}$$

Now Eq. (7.73) is an eigen value equation in SL form, for which we identify

$$p(\rho) = \rho^2, \quad q(\rho) = -l(l+1), \quad \lambda = \frac{(z_{l,m})^2}{a^2}, \quad \text{and} \quad w(\rho) = \rho^2.$$

Hence, $p(\rho) = 0$ gives $\rho = 0$ as one of the extremities of the interval. For the other (corresponding to physically restricted boundary), we choose $\rho = a$ (so that mathematical interval becomes $0 \le \rho \le a$), together with the additional boundary condition for hermiticity (Dirichlet boundary condition)

$$j_l\left(\frac{z_{l,m}}{a}\rho\right)\Big|_{\rho=a} = j_l\left(z_{l,m}\right) = 0. \tag{7.74}$$

Thus $z_{l,m}$ is the m-th zero of $j_l(x)$. With this condition, Eq. (7.73) is an eigen value equation with its differential operator becoming Hermitian in the interval $[0, a]$. The eigen functions are orthogonal in this mathematical interval w.r.t. the weight function. The complete orthonormality relation is (see Eq. (5.99) of Chattopadhyay 2006)

$$\int_0^a j_l\left(\frac{z_{l,m}}{a}\rho\right) j_l\left(\frac{z_{l,n}}{a}\rho\right)\rho^2 d\rho = \frac{a^3}{2}\left[j_{l+1}\left(z_{l,m}\right)\right]^2 \delta_{m,n}. \tag{7.75}$$

Note that the factor ρ^2 in the integrand comes from the weight function and the mathematical interval is $0 \le \rho \le a$.

10. **Spherical modified Bessel equation**

$$x^2 y''(x) + 2x\, y'(x) - \left[x^2 + l(l+1)\right] y(x) = 0. \qquad (l \text{ is an integral constant}) \tag{7.76}$$

This has the same singularities as the other Bessel equations and appears in problems with spherical symmetry, which includes absorption or decay. Here x is the radial distance of a point from the origin. Substitution of $y(x) = \frac{Z_{l+\frac{1}{2}}(x)}{\sqrt{x}}$ in Eq. (7.76) leads to Eq. (7.64) with $\nu = l + \frac{1}{2}$. Hence its solutions called spherical modified Bessel function of first and second kind are defined as (Arfken 1966)

$$i_l(x) = \sqrt{\frac{\pi}{2x}}\, I_{l+\frac{1}{2}}(x)$$
$$k_l(x) = \sqrt{\frac{2}{\pi x}}\, K_{l+\frac{1}{2}}(x), \tag{7.77}$$

where the constants are arbitrarily chosen for convenience. One can verify that $i_l(x)$ is finite at $x = 0$ and diverges for large x, while $k_l(x)$ vanishes in the limit $x \to \infty$ and diverges at $x = 0$. These functions are not oscillatory, but monotonic. See Arfken (1966) for plots of $i_0(x)$, $i_1(x)$, $k_0(x)$ and $k_1(x)$. A general solution of Eq. (7.76) is a linear combination of $i_l(x)$ and $k_l(x)$, satisfying the boundary conditions.
Recurrence relation satisfied by $i_l(x)$ is (Abramowitz 1972)

$$i_{l-1}(x) - i_{l+1}(x) = \frac{(2l+1)}{x}\, i_l(x),$$
$$l\, i_{l-1}(x) + (l+1)\, i_{l+1}(x) = (2l+1)\, i_l'(x), \tag{7.78}$$

and a similar one for $k_l(x)$.
Like modified Bessel functions, spherical modified Bessel functions also do not satisfy orthogonality relations. See the discussion on existence of orthogonality for the former functions, which applies in the present case also.

7.3 Problems

1. Show that two functions $y_1(x)$ and $y_2(x)$ are linearly independent if they are not proportional to each other.
2. Study the nature of singularity for $x \to \pm\infty$ of the equation $y''(x) + \omega^2 y(x) = 0$.
3. Show that the value of index k is modified if we do not impose the condition $a_0 \neq 0$ in Eq. (7.10), but the solution remains essentially the same.
4. Show that we do not get any new linearly independent solution if we choose $a_1 \neq 0$ in Eq. (7.17).
5. Prove Eqs. (7.16) and (7.18) by verifying the general recursion relations leading to these solutions, using mathematical induction.
6. For the following second order differential equations locate the singularities and find their nature. Next obtain the indicial equation and investigate the possibility of obtaining both the linearly independent solutions about $x = 0$ by Frobenius method:

   (a) $y''(x) - \frac{b(b+1)}{x^2} y(x) = 0,\ b > 0)$.
   (b) $y''(x) - \frac{b(b+1)}{x^3} y(x) = 0,\ (b > 0)$.
   (c) $y''(x) + \frac{1}{x} y'(x) - \frac{b^2}{x^2} y(x) = 0,\ (b > 0)$.

(d) $y''(x) + \frac{1}{x^2}y'(x) - \frac{b^2}{x^2}y(x) = 0,\ (b > 0)$.

In case the solution is an infinite series, find if it is convergent. If not, find if it can be truncated to a polynomial by imposing a condition on the parameter b. Discuss the singularity of each one of the solutions.

7. Consider a differential equation

$$\hat{\mathcal{L}}(x)u(x) = p(x)u''(x) + r(x)u'(x) + q(x)u(x) = 0.$$

Find a function $f(x)$ which puts the above equation in SL form, when the equation is pre-multiplied with $f(x)$.

8. Locate all the singularities and find their nature for the following differential equations. Then put them in SL form.

(a) Hermite equation.
(b) Laguerre equation.
(c) Bessel equation.
(d) Hypergeometric equation:
$x(1-x)\,y''(x) + [c - (a+b+1)x]\,y'(x) - ab\,y(x) = 0,$
where a, b and c are constants.
(e) Confluent hypergeometric equation:
$x\,y''(x) + (c-x)\,y'(x) - a\,y(x) = 0,$
where a and c are constants.

References

Abramowitz, M., Stegun, I.A. (eds.): Handbook of Mathematical Functions. Dover Publications Inc., New York (1972)

Arfken, G.: Mathematical Methods for Physicists. Academic Press, New York (1966)

Bell, W.W.: Special Functions for Scientists and Engineers. D. Van Nostrand Co., Ltd., London (1968)

Chattopadhyay, P.K.: Mathematical Physics. New Age International (P) Ltd., New Delhi (2006)

Chapter 8
Solution of Schrödinger Equation: Boundary and Continuity Conditions in Coordinate Representation

Abstract In this chapter, it is shown in general how boundary and continuity conditions lead to eigen solutions: discrete for localized and continuous for non-localized systems. Bound, unbound and quasi-bound systems are discussed. Introducing quantum numbers, importance of symmetry in choosing coordinate system and its connection with degeneracy have also been discussed. Wave packets, Ehrenfest's theorem and their relation with classical physics have been presented.

Keywords Boundary and continuity conditions on wave function · Proof that these lead to eigen solutions · Properties of eigen solutions · Motion of wave packets · Ehrenfest's theorem

In Chap. 9 we will discuss the solution of the time-independent differential Schrödinger equation in one dimension, Eq. (7.3), for specific real potentials. Before that, in this chapter we discuss the conditions that must be satisfied and the general features of the time-independent solutions of the one-dimensional Schrödinger equation (7.3). Three-dimensional cases are straight forward generalization of the one-dimensional case.

8.1 Conditions on Wave Function

The applied force field (produced by the potential) can restrict the motion of a particle to a limited region of space or it may permit the particle to be found at infinity. The state of the particle in the former case is called a *localized* state and in the latter a *scattering* state. The localized state can again be classified into two: a bound state (when the particle is localized within a *fixed region of space*) and a *traveling wave packet* (when the particle can be found within a finite region of space which moves with time). A traveling wave packet can also be scattered by a force field. All these states satisfy different boundary conditions, which we discuss in the following.

Boundary conditions
$|\psi(\vec{r}, t)|^2$ is the position probability density which is a measurable quantity. Hence it must be finite and continuous everywhere. Consequently, $\psi(\vec{r}, t)$ *must be finite and continuous everywhere*. We have the following situations

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5_8

1. For a particle which is *bound within a specified region of space* due to confinement by a potential (a bound state problem), the probability of finding the particle at large distances outside this region is zero, Hence $\psi(\vec{r}, t) \to 0$ as $r \to \infty$.
2. If the particle is represented by a traveling finite *wave packet*, then at any time it is restricted in space (i.e. the probability of finding it vanishes outside a finite region of space, the region shifting with time) and so $\psi(\vec{r}, t) \to 0$ as $r \to \infty$.

   In both these situations, the normalization is done by requiring that ψ represents a single particle, so that the total probability of finding the particle anywhere in space is unity

$$\int |\psi(\vec{r}, t)|^2 d^3r = 1, \tag{8.1}$$

   where $\int$ means integration over all space, unless otherwise specified. Since $\psi(\vec{r}, t)$ vanishes at infinity, the discussion following Eq. (3.30) shows that the normalization will be maintained at all times.
3. In a simple *scattering problem* a particle comes from a large distance, interacts with a potential field $V(\vec{r})$ and flies again to a great distance. In this case there is a finite probability of finding the particle at large distances, for the time-independent solution. So the wave function *does not vanish* at infinity, but it *must remain finite* as $r \to \infty$. In this case, the normalization integral (8.1) diverges and the normalization of the wave function is done by one of the three possibilities: *flux normalization*, *box normalization* or *delta function normalization* (see Sect. 3.5, Chap. 3). It should be noted that the final result is independent of the normalization chosen.
   Appropriate changes are to be incorporated for scattering of a traveling wave packet by a fixed potential. We will not discuss such scattering in this book.

Continuity conditions

In order that physical quantities, *viz.*, position probability density, $\rho(\vec{r}, t) = \psi^*(\vec{r}, t)\,\psi(\vec{r}, t)$, and momentum density $\psi^*(\vec{r}, t)\left(-i\hbar\vec{\nabla}\right)\psi(\vec{r}, t)$ (or in case of a traveling wave or a wave packet, probability current density, $\vec{j}(\vec{r}, t) = \frac{\hbar}{2im}\left[\psi^*(\vec{r}, t)\vec{\nabla}\psi(\vec{r}, t) - \left(\vec{\nabla}\psi^*(\vec{r}, t)\right)\psi(\vec{r}, t)\right]$) are finite, continuous and single valued (as required by experience), both $\psi(\vec{r}, t)$ and $\vec{\nabla}\psi(\vec{r}, t)$ must be *finite, continuous and single valued* at every point in space, so that a definite physical situation can be represented uniquely by a wave function. This condition is sometimes stated as *the wave function must be well behaved.*

Connection between physical domain and mathematical singularities of the second order differential equation

The wave function is the solution of the differential Schrödinger equation (subject to boundary conditions). The condition of being *well behaved* requires the domain to be contiguous and free from singularities. We have seen in Chap. 7 that this condition is always satisfied by second order differential equations encountered in quantum mechanics. The *mathematical domain* is defined by choosing the location

of singularities as one or both the boundaries. This agrees with the *natural (physical) domain*, unless it is limited by imposed restrictions on the motion of the particle (e.g., a particle confined in a spherical hole, restricting its possibility to go to large distances).

A solution of the second order differential equation involves two arbitrary constants. These constants and the energy eigen value are determined by the boundary and continuity conditions, as we will see in Chap. 9, for specific cases. In the following we will see (following Schiff 1968) in general how these conditions arise.

Boundary conditions at an infinite potential step

The natural interval gets limited when a particle encounters a rigid (i.e. impenetrable) wall. In this case, the potential jumps to infinity discontinuously. This restricts the physical interval and the boundary condition must come from physical considerations.

Suppose $V(\vec{r})$ has an infinite discontinuity across a continuous surface, so that V is finite (we take it zero) on one side and $+\infty$ on the other. We consider a tangent plane to the surface at the point of interest and a line perpendicular to this plane at that point is taken as the x-axis. Then the potential is discontinuous along x, but continuous along the y- and z-axes.[1] This becomes an effectively one-dimensional problem with

$$\begin{aligned} V(x) &= 0 \quad \text{for } x < 0, \\ &= \infty \quad \text{for } x \geq 0. \end{aligned}$$

To solve the Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2u(x)}{\mathrm{d}x^2} + V(x)u(x) = Eu(x), \tag{8.2}$$

we first consider the wall to have a finite height V_0

$$\begin{aligned} V(x) &= 0 \quad \text{for } x < 0, \\ &= V_0 \quad \text{for } x \geq 0, \end{aligned} \tag{8.3}$$

and then take the limit $V_0 \to \infty$. Hence we take $E < V_0$. We name the regions $x < 0$ and $x \geq 0$ as regions I and II respectively (see Fig. 8.1).

In region I

$$\frac{\mathrm{d}^2u_I(x)}{\mathrm{d}x^2} + \alpha^2 u_I(x) = 0 \quad \text{with } \alpha = \sqrt{\frac{2m}{\hbar^2}E},$$

choosing the positive root, $\alpha > 0$ (a negative root will simply interchange the sign of arbitrary constant A below). A general solution is

[1] Actually, the infinite jump in potential is a mathematical idealization. In reality, the potential can be very large but finite. This is akin to the limiting process adopted below.

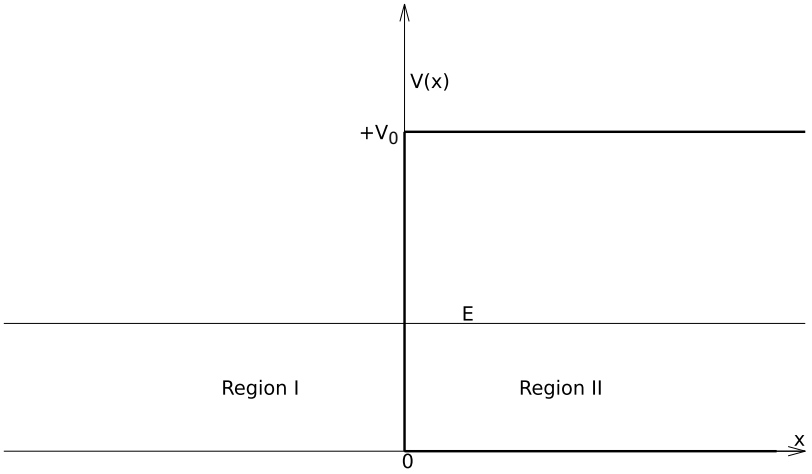


Fig. 8.1 Regions I and II for the potential $V(x)$

$$u_{I}(x) = A \sin \alpha x + B \cos \alpha x. \tag{8.4}$$

In region II, we have from Eqs. (8.2) and (8.3)

$$\frac{\mathrm{d}^2 u_{II}(x)}{\mathrm{d}x^2} - \beta^2 u_{II}(x) = 0 \quad \text{with } \beta = \sqrt{\frac{2m}{\hbar^2}(V_0 - E)},$$

choosing the positive root, $\beta > 0$ (a negative root will simply interchange the arbitrary constants C and D below). A general solution of this equation is

$$u_{II}(x) = C\mathrm{e}^{-\beta x} + D\mathrm{e}^{\beta x}.$$

By the condition that $u_{II}(x)$ must be finite as $x \to \infty$, we must have

$$D = 0,$$

and the continuity of wave function at $x = 0$ gives from Eq. (8.4)

$$B = C.$$

Hence

$$u_{I}(x) = A \sin \alpha x + C \cos \alpha x,$$
$$u_{II}(x) = C\mathrm{e}^{-\beta x}.$$

By the condition of the continuity of derivatives

$$\frac{\mathrm{d}u_I(x)}{\mathrm{d}x}\Big|_{x=0} = \frac{\mathrm{d}u_{II}(x)}{\mathrm{d}x}\Big|_{x=0}$$

we have

$$A\alpha = -C\beta. \tag{8.5}$$

Now let $V_0 \to \infty$. Then $\beta \to \infty$. Left side of Eq. (8.5) must remain finite (as A is finite for the wave function $u_I(x)$ must be finite and E is finite). Hence in the limit $\beta \to \infty$, we must have $C \to 0$. Thus in the limit $V_0 \to \infty$, we have

$$u_I(x) = A \sin \alpha x,$$
$$u_{II}(x) = 0.$$

Note that $-C\beta$ becomes indeterminate in this limit. Hence the boundary conditions at an infinite potential step are

1. Wave function must vanish at the point of infinite discontinuity and also vanishes in the region where $V \to \infty$. This is so, because for $u(x) \neq 0$, the term $V(x)u(x)$ diverges and Eq. (8.2) cannot be satisfied.
2. Gradient of the wave function normal to the surface of discontinuity is indeterminate and is not continuous at the point of infinite discontinuity. Since the gradient is discontinuous, momentum is also discontinuous. This can be understood as follows. The particle cannot penetrate and gets reflected by the infinite wall. Hence momentum for $x > 0$ is zero, while it is finite for $x < 0$. This is so even classically, for which the infinitely rigid wall absorbs the change in momentum at $x = 0$ of the particle traveling in positive and negative x-directions, as it gets reflected by the wall.

8.2 Eigen Solutions

Boundary conditions lead to energy eigen values: one-dimensional qualitative treatment

We will see that the application of boundary and continuity conditions will lead to energy *eigen values and eigen functions*, which are the only allowed non-trivial solutions of Eq. (8.2). For any other value of E, the only quantum mechanically acceptable solution will be the trivial solution $u(x) = 0$.

Following Schiff (1968), we will consider a general one-dimensional potential shown in Fig. 8.2. For simplicity we take $V(x)$ symmetrical about $x = 0$, with a minimum $V_{\min}$ at $x = 0$ and it vanishes asymptotically, $V(x) \to 0$ as $|x| \to \infty$. Energy E is shown as a horizontal line. Classically the particle cannot exist for regions where $V(x) > E$, since then kinetic energy $(E - V(x)) < 0$ and velocity

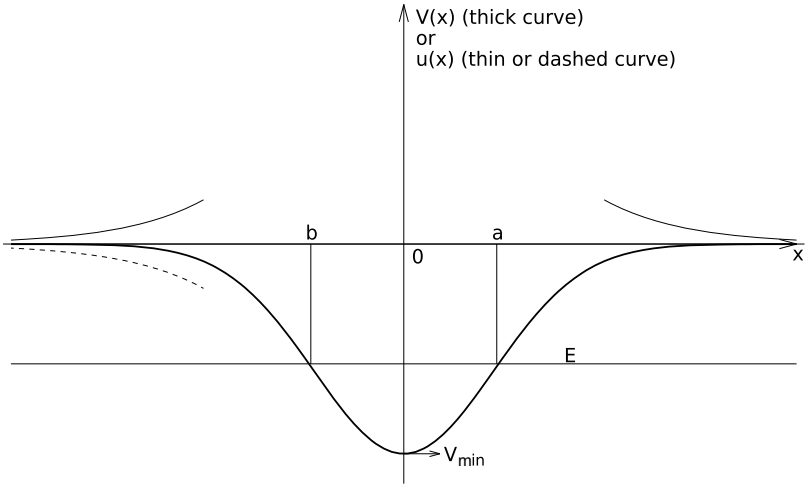


Fig. 8.2 A general one-dimensional potential well together with $u(x)$ in the asymptotic regions (with a large multiplying factor, so as to vividly show its mathematical nature)

becomes imaginary. The points where $V(x) = E$ (shown as 'a' and 'b' in Fig. 8.2) are called *classical turning points* (CTP), since the classical motion is restricted to $b \leq x \leq a$ and is turned around, i.e. reversed at points $x = a$ and $x = b$.

We first consider $E < 0$, so that $E = -B$ with $B > 0$. B is called the *binding energy*. It is the energy needed by the particle in order to be free, i.e. to be outside the potential well. In the asymptotic region ($|x| \to \infty$), Eq. (8.2) becomes

$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} - \beta^2 u = 0, \quad \text{where } \beta = \sqrt{\frac{2m}{\hbar^2} B}, \qquad \text{(choose } \beta \text{ positive)}.$$

whose general solution is

$$u(x) = C\mathrm{e}^{-\beta x} + D\mathrm{e}^{\beta x}.$$

(Note that if β is chosen negative, arbitrary constants C and D will simply get interchanged). Now $u(x)$ must be finite everywhere. Since for $x \to +\infty$, the second term diverges, D must vanish for $x \to \infty$. Similarly for $x \to -\infty$, C must vanish. Denoting u by $u_<$ and $u_>$ for regions $x < 0$ and $x > 0$ respectively

$$\begin{aligned} u_<(x) &\propto \mathrm{e}^{\beta x} \quad \text{for } x \to -\infty, \\ u_>(x) &\propto \mathrm{e}^{-\beta x} \quad \text{for } x \to \infty. \end{aligned} \tag{8.6}$$

Thus the asymptotic boundary conditions determine the nature of asymptotic wave functions.

Next we see the effect of the continuity conditions of $u(x)$ and $\frac{\mathrm{d}u}{\mathrm{d}x}$ throughout the domain of x. To eliminate the overall normalization constant, we consider the

continuity of $\frac{1}{u}\frac{\mathrm{d}u}{\mathrm{d}x}$. Since a specific differential equation (without a singularity in the entire domain) is solved in each of the three regions $-\infty < x \leq b$, $b \leq x \leq a$ and $a \leq x < \infty$, the quantity $\frac{1}{u}\frac{\mathrm{d}u}{\mathrm{d}x}$ within each region is continuous. Hence this condition should be imposed at $x = a$ only (due to symmetry the condition at $x = b$ is automatically satisfied). Although the sign of u is arbitrary, the quantity $\frac{1}{u}\frac{\mathrm{d}u}{\mathrm{d}x}$ is independent of this arbitrariness. From Eq. (8.6), we see that this quantity will be positive for $x \to -\infty$ and negative for $x \to \infty$. With chosen initial values of u and $\frac{\mathrm{d}u}{\mathrm{d}x}$ at far-off starting points, one can integrate Eq. (8.2) inwards[2] from the asymptotic regions to get these at $x = 0$. Whether $\frac{1}{u}\frac{\mathrm{d}u}{\mathrm{d}x}$ obtained in this manner from the two asymptotic regions match at $x = 0$ for any value of E will be seen in the following.

Rewriting Schrödinger equation as

$$\frac{1}{u}\frac{\mathrm{d}^2u}{\mathrm{d}x^2} = \frac{2m}{\hbar^2}\big(V(x) - E\big), \tag{8.7}$$

we see that for the non-classical regions ($E < V(x)$)

$$\frac{1}{u}\frac{\mathrm{d}^2u}{\mathrm{d}x^2} > 0. \tag{8.8}$$

Hence the curve $u(x)$ is convex toward the x-axis. This is seen to be true in Fig. 8.2 for the asymptotic regions. A negative sign of $u(x)$ does not change the convex nature, as can be seen by the dashed curve for $x \to -\infty$. In the following, we consider increasing E gradually, starting from a value below $V_{\min}$.

Case 1. $E < V_{\min}$

Inequality (8.8) shows that $u(x)$ remains convex toward the x-axis in the entire region, both for $x < 0$ and $x > 0$ (see Fig. 8.3). We can easily convince ourselves that for a curve to be convex in $x < 0$ region, both $u(x)$ and $\frac{\mathrm{d}u(x)}{\mathrm{d}x}$ are positive (continuous curve in Fig. 8.2) or both are negative (dashed curve in Fig. 8.2). This is clearly true in the asymptotic region $x \to -\infty$ (Fig. 8.2), as seen from Eq. (8.6). The solutions are then extended inwards (say, by point-to-point integration) up to the origin. Since the curve is convex, we can see from Fig. 8.3 that for $x > 0$, $u(x)$ and $\frac{\mathrm{d}u(x)}{\mathrm{d}x}$ are of opposite signs. Hence a quantity $\mathcal{R}(E, x) \equiv \frac{1}{u}\frac{\mathrm{d}u}{\mathrm{d}x}$ will remain positive for the entire region $x < 0$ and it will remain negative for the entire region $x > 0$. The two solutions cannot match at $x = 0$ as shown in Fig. 8.3. If the wave functions are matched by adjusting the normalization constants (continuous curves), their slopes do not match. On the other hand, if slopes are matched (dashed curve for $x < 0$ and continuous curve for $x > 0$), the wave function becomes discontinuous. Hence continuity of both $u(x)$ and $\frac{\mathrm{d}u}{\mathrm{d}x}$ cannot be satisfied. Thus there is no acceptable solution for any $E < V_{\min}$. This agrees with the classical situation.

[2] We can visualize the process of integration as point-to-point integration in small steps, adopted in numerical integration.

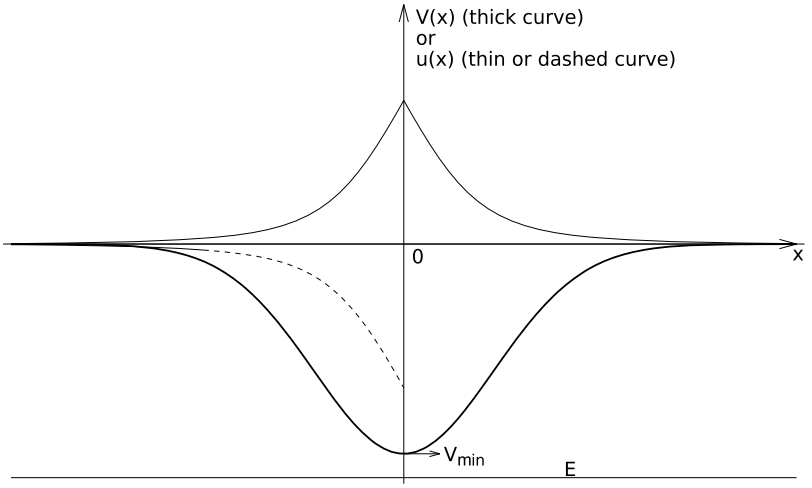


Fig. 8.3 Plot of $u(x)$ for an energy $E < V_{\min}$, for both $x < 0$ and $x > 0$ (thin continuous curves). The dashed curve shows a plot of $-u(x)$ for $x < 0$ only

Case 2. $V_{\min} < E < 0$

In this case two turning points appear at 'a' and 'b' (Fig. 8.4). For the region $b < x < a$, where $E > V(x)$, Eq. (8.7) shows

$$\frac{1}{u}\frac{\mathrm{d}^2u}{\mathrm{d}x^2} < 0.$$

Thus $u(x)$ curve is *concave* toward x-axis inside the region $b < x < a$ and *convex* outside it. Although $\mathcal{R}(E, x) = \frac{1}{u}\frac{\mathrm{d}u}{\mathrm{d}x}$ still remains positive for $x < 0$ and negative for $x > 0$, their magnitudes decrease in $b < x < a$ as one approaches the origin, because of the concave curvature of $u(x)$. Note that at the CTP, $\frac{\mathrm{d}^2u}{\mathrm{d}x^2} = 0$ and $\frac{\mathrm{d}u}{\mathrm{d}x}$ has an extremum. Note also that due to the symmetry of the potential, the magnitudes of $\mathcal{R}(E, x)$ from the two sides will be same, but of opposite signs at the origin. For E just above $V_{\min}$, the quantities $\mathcal{R}(E, x)$ for $x < 0$ and $x > 0$ will *not* match at $x = 0$, since the concave region is not large enough.

By adjusting the constants, if $u(x)$ is matched, then $\frac{\mathrm{d}u}{\mathrm{d}x}$ becomes discontinuous (continuous curves in Fig. 8.4). If on the other hand $\frac{\mathrm{d}u}{\mathrm{d}x}$ is matched (by changing the sign of $u(x)$ for $x < 0$ region), $u(x)$ becomes discontinuous (dashed curve for $x < 0$ and continuous curve for $x > 0$ in Fig. 8.4). Since $\mathcal{R}(E, x)$ is not continuous at $x = 0$, there is no acceptable solution for this E.

However as E increases (still negative), the turning points are further apart, making the concave region wider and reducing $|\mathcal{R}(E, x)|$, as one approaches the origin from either side. Ultimately, we will find a value of E ($= E_0$, still negative), for which $\mathcal{R}(E_0, 0) = 0$. Then $\mathcal{R}(E_0, 0)$ from $x < 0$ and $x > 0$ will match. This is shown by the continuous curve in Fig. 8.5, in which $u(x)$ matches (by adjusting arbitrary

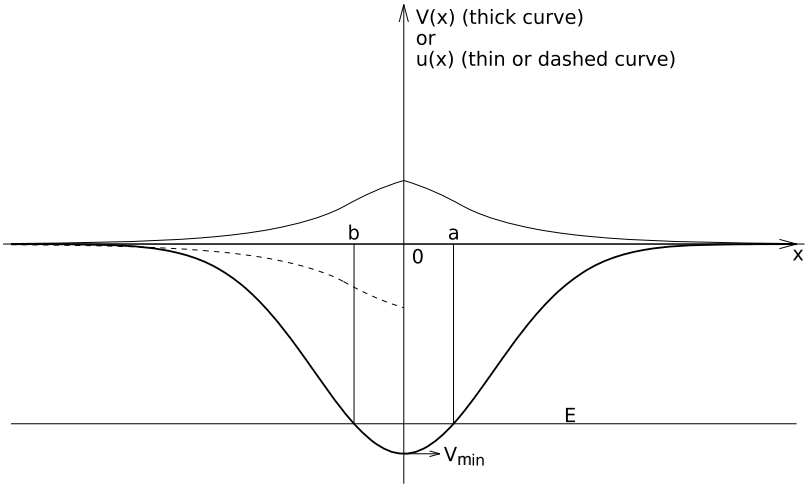


Fig. 8.4 Plot of $u(x)$ for an energy slightly larger than $V_{\min}$ for both $x < 0$ and $x > 0$ (thin continuous curves). The dashed curve shows a plot of $-u(x)$ for $x < 0$ only

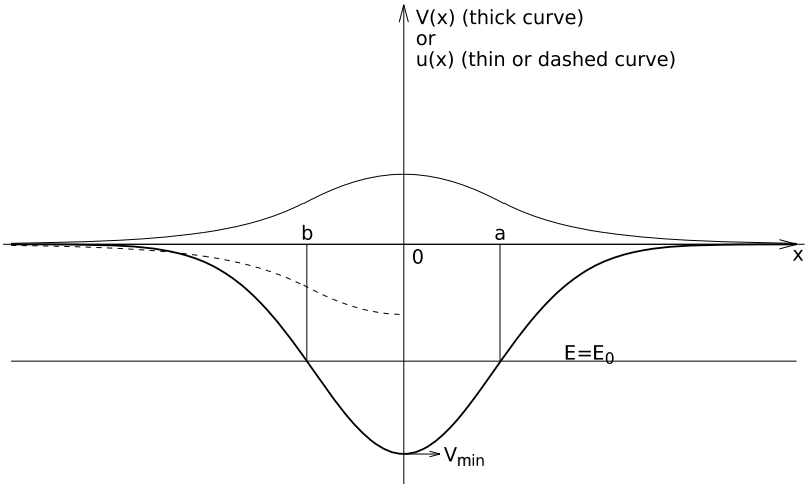


Fig. 8.5 Plot of $u(x)$ for $x < 0$ and $x > 0$ (continuous thin curves) corresponding to an energy $E = E_0$, for which both u and $\frac{\mathrm{d}u}{\mathrm{d}x}$ match. This is the lowest lying acceptable solution, called ground state (the dashed curve shows a plot of $-u(x)$ for $x < 0$ only)

constants in asymptotic regions) and $\frac{\mathrm{d}u}{\mathrm{d}x}$ also matches, since it is zero (from both sides) at the origin. Thus E_0 is an allowed energy. It is the lowest allowed energy and is called the *ground state energy*; the corresponding $u(x)$ is the ground state wave function (denoted by $u_0(x)$). Note that it has no nodes.

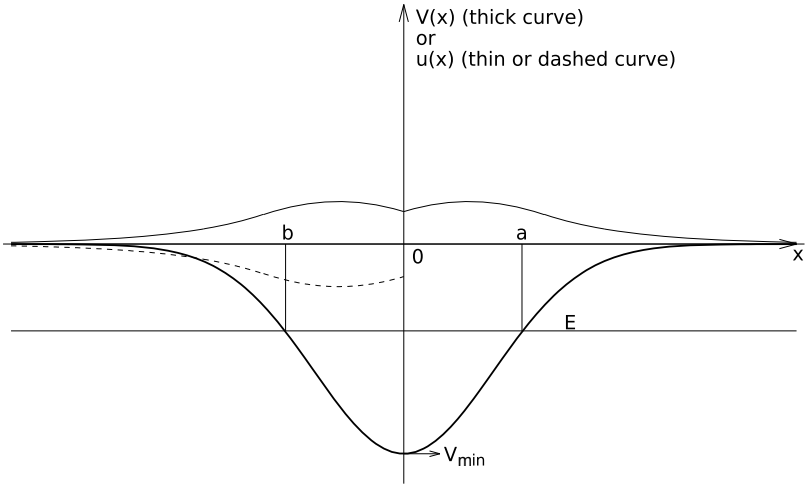


Fig. 8.6 Plot of $u(x)$ corresponding to an energy E slightly greater than E_0. Either $\frac{du}{dx}$ is discontinuous (continuous curves) or u is discontinuous (dashed curve for $x < 0$ and continuous for $x > 0$). Hence it is not an acceptable solution

We can easily convince ourselves that for E slightly larger than E_0, again $\mathcal{R}(E, x)$ from the two sides will not match at the origin (see Fig. 8.6).

However if the potential is deep and wide enough, we may find another negative energy $E = E_1$ with $E_0 < E_1 < 0$, for which $u_<(0) = 0 = u_>(0)$. But the corresponding first derivatives are discontinuous, having same magnitude, but with opposite signs at the origin (dashed curve for $x < 0$ and continuous curve for $x > 0$ in Fig. 8.7). However, we can choose the constant for $u_<(x)$ negative, such that $u_<(x)$ is negative, while $u_>(x)$ is positive. Then both u and $\frac{du}{dx}$ from $x < 0$ and $x > 0$ match at the origin. With this, $u_<(x)$ (negative curve) for $x < 0$ and $u_>(x)$ (positive curve) for $x > 0$ are continuous, as also are their first derivatives, at the origin. Hence this energy is acceptable, as also the corresponding wave function (continuous curve in Fig. 8.7). Thus we can have another energy eigen value (E_1) and eigen function $u_1(x)$ corresponding to the *first excited state*. There is one node in $u_1(x)$, which is at the origin for the symmetric potential.

We can see that for E slightly greater than E_1 (but still negative), the curves for $x < 0$ and $x > 0$ cannot be matched (see Fig. 8.8).

We can continue in this manner and see that if the potential is deep and wide enough, we may get more allowed energy eigen values corresponding to second, third, $\cdots$ excited states. Figure 8.9 shows the second excited state for $E = E_2$ (still negative) for which $\frac{du}{dx} = 0$ from both sides and $u(x)$ is matched at the origin. The corresponding eigen function is $u_2(x)$, which has two nodes.

For the chosen symmetric but otherwise arbitrary potential, it is seen that $\frac{du}{dx}|_{x=0} = 0$ for the ground, second excited, fourth excited, etc. states, while $u(0) = 0$ for the first excited, third excited, etc. states. This is because the wave function is symmetric

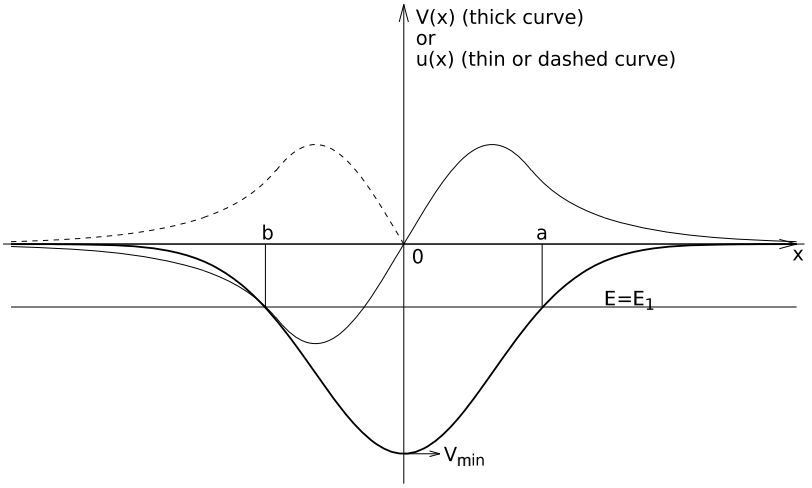


Fig. 8.7 Plot of $u(x)$ corresponding to an energy $E = E_1$ for which $u(0) = 0$ from both sides. The continuous curve is the wave function of the first excited state

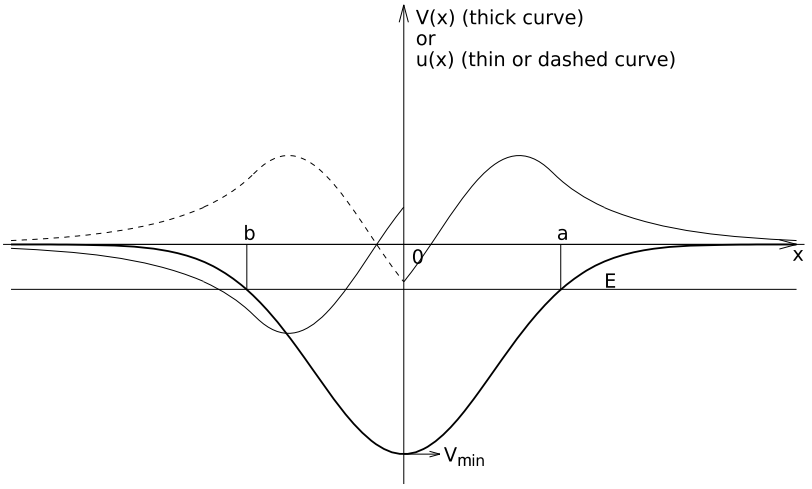


Fig. 8.8 Plot of $u(x)$ corresponding to an energy E slightly greater than E_1. This is not an eigen function, since both u and $\frac{\mathrm{d}u}{\mathrm{d}x}$ cannot be matched

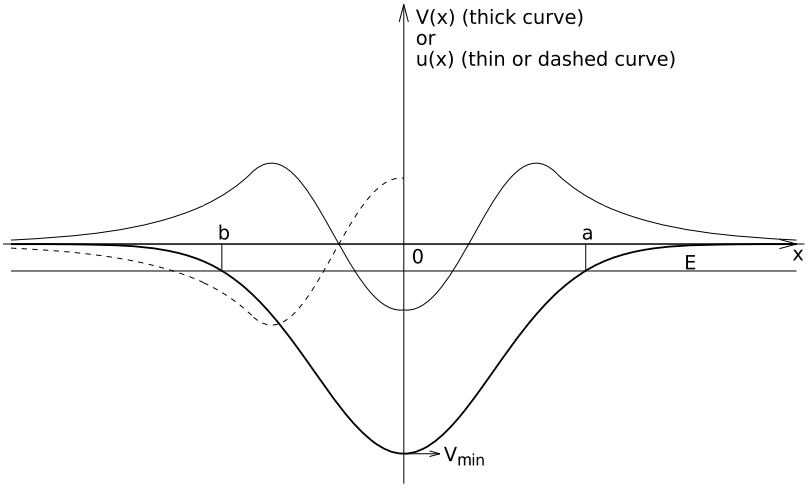


Fig. 8.9 Plot of $u_2(x)$ for the second excited state (continuous curve). Other curves are as in previous figures

for the first group of states, while it is antisymmetric for the second group. This is due to the fact that the potential is symmetric about $x = 0$ (see later for discussion of parity).

Incidentally we see that the ground state has no nodes, the first excited state has one node, and so on. The n-th excited state has n nodes ($n = 1, 2, \cdots$).

Case 3. $E > 0$

In the asymptotic region $V(x) = 0$ and the Schrödinger equation (8.2) for $E > 0$ is

$$\frac{\mathrm{d}^2u}{\mathrm{d}x^2} + \alpha^2 u = 0, \quad \text{where } \alpha = \sqrt{\frac{2m}{\hbar^2}E} \qquad \text{with } \alpha \text{ positive,}$$

whose solution is

$$u(x) = A \sin \alpha x + C \cos \alpha x.$$

Since both $\sin \alpha x$ and $\cos \alpha x$ remain finite even for $|x| \to \infty$, both these can survive, unlike the case for $E < 0$ when only the outwardly decreasing exponential can survive. Hence we have two arbitrary constants (A and C) in both $x < 0$ and $x > 0$ regions. We can always match both $u(x)$ and $\frac{\mathrm{d}u(x)}{\mathrm{d}x}$ arising from the left ($x < 0$) and the right ($x > 0$) solutions, by adjusting two arbitrary constants in both regions. Suppose the two constants are chosen for $x \to -\infty$ region and Eq. (8.2) integrated from a large negative x to $x = 0$. With this fixed, Eq. (8.2) can again be integrated backwards from a large positive x to $x = 0$ and the two constants in this region can be adjusted to make both $u(x)$ and $\frac{\mathrm{d}u}{\mathrm{d}x}$ continuous at $x = 0$ (two equations for two unknowns). This can be done for *any positive value of E*. Hence all possible values of

$E > 0$ are allowed. Thus we have continuous energy eigen values. But such solutions do not represent bound states, since there is a finite probability of finding the particle as $|x| \to \infty$.

General features

From a generalization of the above features, we have the following:

1. *A bound state cannot have energy equal to that of the bottom of a well.*
   We can see this from Fig. 8.3: if $E = V_{\min}$ then $u(x)$ will be *convex toward x-axis for both* $x < 0$ *and* $x > 0$ and both increase as they approach the origin. Hence the wave functions in the two regions cannot be matched. Thus the system must have a minimum energy above the bottom of the well (called the *zero point energy*) to exist, in stark contrast with the classical situation of rest. Classically the particle can be at rest at $x = 0$ with $E_{cl} = V_{\min}$, hence $p_{cl} = 0$. Thus both position and momentum are specified simultaneously. This is possible in classical mechanics, but not in quantum mechanics, for which the uncertainty principle requires the system to have a zero point motion (as against being completely at rest), even in the ground state.
2. If $V(x)$ is finite negative in some region, with a minimum $V_{\min}$, and goes to zero as $|x| \to \infty$, then there will be one or more discrete energy eigen values with $V_{\min} < E < 0$. There will be *at least one bound state*, no matter how shallow or narrow the one-dimensional potential is (see below and Problem 1). However for a three-dimensional potential, a minimum well depth is necessary for the existence of even the ground state (see below and Sect. 10.3 of Chap. 10).
   If $V(x)$ does not vanish for $|x| \to \infty$ but has a minimum $V_{\min}$ somewhere, then there will be discrete eigen values within the energy interval $V_{\min} < E < V_M$, where V_M is the smaller of $V(-\infty)$ and $V(+\infty)$ (see Problem 2).
3. For the last case, if E is larger than V_M, there will be continuous energy eigen values, as the probability of finding the particle in the asymptotic region with the smaller wall is non-vanishing (Problem 2).
4. The lowest energy eigen function (called *the ground state*) has no nodes, the next allowed one (called *the first excited state*) has one node, the next one (called *the second excited state*) has two nodes, and so on. The n-th excited state has n nodes.
5. A discrete energy eigen function is *localized*; i.e. it vanishes as $x \to \pm\infty$. Since the probability of finding the particle vanishes outside a certain region, such a state is called a *bound state*.
6. From the above discussion, it is intuitively clear that the number of discrete energy bound eigen states for a given potential will increase with the 'depth' (V_d) and 'width' (a_w) of the potential. The depth and width are well defined for a square well potential, but not for an arbitrary potential well. A crude measure of depth can be given in terms of $V_M - V_{\min}$. Similarly a crude measure of width can be given in terms of the extent of x, over which bound states are possible. For a finite square well of width $2a$ and depth V_0, we will see in Sect. 9.2 of Chap. 9, that the number of eigen states increases with $V_0 a^2$. For more discussion; see Sect. 8.3 of this chapter.

7. We already stated that for a one-dimensional potential well (with positive V_d and a_w) at least one bound state is always possible. As an example one can see this from Fig. 9.3 for a finite square well, since one solution is always possible even for very small $V_0 a^2$. For the general case of a very shallow well, we can see this from Fig. 8.4. For very small negative E, the curvature of the convex part can be made as small as we wish. Then even a small concave part can bring $\frac{\mathrm{d}u}{\mathrm{d}x}|_{x=0}$ to zero, leading to the ground state (Problem 1).
8. A particle can be temporarily trapped in a suitable potential. As a simple example, consider the one-dimensional potential of Fig. 8.10, which is ∞ for $x < 0$, is zero in $0 \leq x \leq a$, then a positive finite barrier of height V_0 in $a < x \leq b$ and vanishes for $x > b$ (where a and b are finite positive constants and $b > a$). The region with $x < 0$ is inaccessible with any energy, as the wave function must vanish in this region for the Schrödinger equation with any finite energy. A particle having energy E ($0 < E < V_0$) will be trapped in Region I, but the wave function in Region III is sinusoidal and does not vanish even as $x \to \infty$. Thus a probability of finding the quantum particle exists even outside the trap. This is called *quantum mechanical tunneling* through the barrier (Region II). A classical particle with E ($< V_0$) in Region I will be permanently trapped. A solution of the Schrödinger equation subject to appropriate boundary conditions shows that its energy is complex (see problems of Chap. 9) with a negative imaginary part, i.e. $E = E_R - iE_I$ with $E_I > 0$ (which is necessary since the total probability of finding the particle over all space, at all times must be finite). This makes the probability density

$$\rho(x,t) = |\psi(x,t)|^2 = |u(x)\mathrm{e}^{-i\frac{(E_R - iE_I)}{\hbar}t}|^2 = |u(x)|^2\mathrm{e}^{-\frac{2E_I}{\hbar}t}$$

finite at all times and its integration over all space remains finite, if $u(x)$ is normalized.[3] However we can see that $\int \rho(x,t)\mathrm{d}x$ will still have a factor $\mathrm{e}^{-\frac{2E_I}{\hbar}t}$, showing that the total probability decreases with time. This corresponds to gradual leaking out of the particle trapped temporarily in Region I. Hence such a state is called a *quasi-bound state*.
For a purely *bound state* E is purely real and $\rho(x,t)$ becomes independent of t, giving rise to a stationary state.
9. For a bound state, the expectation value of momentum vanishes (see this chapter, Sect. 8.6). This can be understood since a bound state is localized in space. Therefore the system does not have a net movement as a whole.

All these features remain valid for the separated radial equation of a spherically symmetric three-dimensional case (see Chap. 10, Sect. 10.3). However a minimum well depth is necessary for existence of the ground state for the three-dimensional case, even for $l = 0$ for which there is no centrifugal repulsion [see Eq. (10.30)]. This can be understood from the following. For a central potential $V(r)$, the radial

[3] Since energy is complex, $u(x)$ will have an exponential damping factor as $x \to \infty$, so that $u(x)$ is normalizable. See discussion following Eq. (9.44).

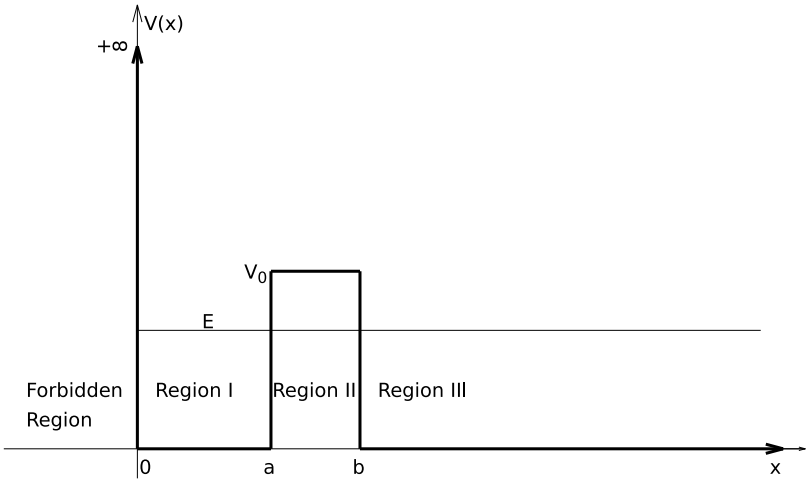


Fig. 8.10 Potential with a finite barrier (thick lines). A particle having energy E $(0 < E < V_0)$ will be temporarily trapped in Region I giving rise to quasi-bound states, which has a probability to leak out (tunnel through) the barrier

equation for $\tilde{R}(r)$ [where $u(\vec{r}) = \frac{\tilde{R}(r)}{r} Y_{lm}(\theta, \phi)$] is similar to the one-dimensional equation with an effective potential $V_{\text{eff}}(r) = V(r) + \frac{\hbar^2}{2m} \frac{l(l+1)}{r^2}$, but the interval of r is $0 \leq r < \infty$ [see Eq. (10.30)]. Since $u(\vec{r})$ must be finite at the origin, $\tilde{R}(0) = 0$. Compare this with the fictitious one-dimensional motion by mentally extending the interval of r to $-\infty < r < \infty$ and reflecting the potential about the origin as $V(|r|)$. Then the effective one-dimensional wave function $\tilde{R}(r)$ appears to have a node at $r = 0$ for this fictitious case. We may then expect a minimum well depth even for $l = 0$, for which $V_{\text{eff}}(r) = V(r)$. However note that it is not a real node of the wave function $u(\vec{r})$, which does not vanish at $r = 0$ for $l = 0$ (see Chap. 11).

8.3 Other Properties

Quantum number

We have just seen that the energy eigen values are *discrete* for bound states. Thus they can be designated by a non-negative integer $(n = 0, 1, 2, \cdots)$ in ascending order, such that E_n is the energy eigen value corresponding to the $(n + 1)$th eigen function. The ground state is denoted by $n = 0$ and has energy eigen value E_0, the nth *excited state* has energy eigen value E_n. The number n is called a *quantum number*. The above discussion shows that the nth excited state has n nodes in the eigen function (for the three-dimensional case, the zero of $\tilde{R}(r)$ at the origin is not counted as a

node; see Sect. 10.3). The quantum number n can thus be taken as the number of nodes in the eigen function.[4]

For bound states in more than one dimension, say in a d-dimensional space, we need d quantum numbers, say, $n_1, n_2, \cdots, n_d$. This can be seen as follows. In a d-dimensional space, we need d variables, which can be chosen in different ways (e.g. Cartesian, polar, etc). Suppose we choose a Cartesian system and consider only the motion in the jth-variable, freezing all remaining $(d-1)$ variables. Then this becomes an effectively one-dimensional problem. The treatment of the last paragraph introduces a quantum number, n_j. This can be repeated for each of the d variables. Thus we need a total of d *independent* quantum numbers. This can also be seen from the Bohr–Sommerfeld quantization rule of the old quantum theory

$$\oint p_j dq_j = n_j h \qquad (j = 1, 2, \cdots, d), \tag{8.9}$$

where p_j is the generalized momentum canonically conjugate to the jth generalized variable q_j and the integral is over a complete cycle of the variable. In a d-dimensional space, we need d generalized variables and for each we need an independent quantum number. Thus in d-dimensional space, we need d independent quantum numbers. Note that this procedure (of old quantum theory) can be applied only to cyclic variables, but gives an idea how a quantum number arises for each variable.

Estimation of the number of bound states in a given potential well

A typical potential well, V, is usually chosen to be negative near origin and vanish asymptotically. Energy eigen values E_n are then negative. However a constant shift (say by C) in V and a consequent shift in E_n by the same amount (C) leave the Schrödinger equation and the eigen functions unchanged. So V may be chosen to be positive and asymptotically equal to C. In this case, E_n is positive and less than C.

We can roughly estimate the number of bound states in a given finite potential well, using the Bohr–Sommerfeld quantization rule, Eq. (8.9). For one-dimension we get

$$\oint p dq = nh = 2\pi\hbar n.$$

We have semiclassically $p = \sqrt{2m(E_n - V(x))}$ for a particle of mass m moving in a potential $V(x)$ with energy E_n. Note that for the maximum value of n $(= n_{\max})$, value of $E_{n_{\max}}$ should be near the top of the finite well, but still less than its asymptotic value. Also, the integral over a complete period ($\oint$) is estimated to be twice the integral over x for its interval. Thus

[4] For H-atom (Chap. 11), the commonly used 'principal quantum number' (n) is actually $n = N + l + 1$, where N is the quantum number for the radial motion (called 'radial quantum number'), such that $N \geq 0$ according to the above discussion. Hence $n \geq 1$. See also discussion following Eq. (10.30).

$$n_{\max} = \text{integer part of} \left[\int_{\text{interval of} x} \sqrt{\frac{2m}{\pi^2\hbar^2}(E_{n_{\max}} - V(x))}\ \mathrm{d}x \right] + 1.$$

Since the integral in the above equation is in general not an integer, $n_{\max}$ is estimated as the integer $\leq$ integral plus 1. As $V(x)$ is commonly taken negative, and it vanishes asymptotically, $E_{n_{\max}}$ should be negative having the smallest magnitude. Hence we can estimate $n_{\max}$ by taking $E_{n_{\max}} \approx 0$ and get

$$n_{\max} = \text{integer part of} \left[\int_{\text{interval of x}} \sqrt{\frac{2m}{\pi^2\hbar^2}(-V(x))}\ \mathrm{d}x \right] + 1. \tag{8.10}$$

We can follow a similar procedure for the radial equation for spherically symmetric potentials in three dimension.

Clearly, for an infinite potential well, there will be an infinite number of bound states. Examples are: infinite square well, (infinite) harmonic oscillator well in one dimension (Chap. 9), H-atom (Chap. 11) and other infinite wells in three dimension (Chap. 12).

Choice of quantum numbers: Conserved observables and good quantum numbers. Importance of symmetry of the Hamiltonian

For a d-dimensional space, we need d quantum numbers. How should we choose them? In one dimension, the only quantum number needed is chosen as the energy eigen value (i.e. eigen value of the Hamiltonian) or more commonly (and equivalently), the sequential cardinal number (n) of the energy eigen value E_n, with n increasing as E_n increases. In more than one dimension, additional quantum numbers are usually related to quantization quantum numbers associated with additional variables of motion, according to Eq. (8.9). For example, for a three-dimensional system having rectangular symmetry (say a particle in a cubical box with rigid walls, see Problems of Chap. 9), quantum numbers can be chosen as $\{n_x, n_y, n_z\}$ associated with variables $\{x, y, z\}$ respectively.

But if we choose the variables $\{x, y, z\}$ in a three-dimensional case with spherical symmetry, the Schrödinger equation does not separate and corresponding quantum numbers $\{n_x, n_y, n_z\}$ are jumbled up in energy and other observables. On the other hand the Schrödinger equation separates nicely in spherical polar variables and it is possible to associate quantum numbers with each variable. Thus one must take the symmetry of the system into consideration (see below). Even in the case of spherical symmetry, all the components of orbital angular momentum $\hat{\vec{L}}$ cannot be specified simultaneously, since they do not commute in pairs (see Chap. 10). The quantum numbers are chosen as $\{n, l, m\}$ corresponding to the Hamiltonian $(\hat{H})$, $\hat{L}^2$ and $\hat{L}_z$ respectively (see Chap. 10). Note that an *eigen value* of one of these operators may be represented by a suitable quantity representing the eigen value. As an example, the eigen value of $\hat{L}^2$ is $l(l+1)\hbar^2$, but it is represented by l only.

To specify an energy level (and the associated state of the system), we need d quantum numbers for a d-dimensional system. These must be specified *simultane-*

ously and are called **good quantum numbers**, if they remain conserved in time. One of these is the Hamiltonian and the remaining $(d-1)$ must correspond to observables each of which must commute with the Hamiltonian [hence these are constants of motion, according to Eq. (4.31)], as well as among themselves in pairs. This set of d observables, consisting of the Hamiltonian and $(d-1)$ others (so that all d operators commute in pairs), is called the **complete set of commuting observables** (CSCO). The CSCO is necessary and sufficient to provide good quantum numbers for classification of a state of the system completely. If some of the $(d-1)$ operators do not commute with the Hamiltonian (but commute among themselves in pairs, so that they can be specified simultaneously and can be used to specify the state of the system), the set of quantum numbers is referred to as *not good*. We discuss some examples below.

As an example consider the H-atom consisting of a proton and an electron, ignoring spin of each particle (see Chap. 11). Only the relative motion, described by the relative vector $\vec{r}$, is of interest. As discussed above, the set of observables $\{\hat{H}, \hat{L}^2, \hat{L}_z\}$ is the CSCO and good quantum numbers are $\{n, l, m\}$. Now suppose we include the spin $(\hat{\vec{S}})$ of the electron into our consideration. This adds two more dimensions (since the components of $\hat{\vec{S}}$ do not commute as those of $\hat{\vec{L}}$). Corresponding observables can be taken as $\hat{S}^2$ and $\hat{S}_z$ with quantum numbers s and m_s respectively. Since s is fixed $(s = \frac{1}{2}\hbar)$, it is usually suppressed. The Hamiltonian remains as before (we rename it $\hat{H}_0$). Then the good quantum numbers are $\{n, l, m_l, m_s\}$ (renaming m by m_l, as in m_s). Next we add the spin-orbit interaction $\hat{\vec{L}}.\hat{\vec{S}}$, so that the Hamiltonian now becomes $\hat{H} = \hat{H}_0 + g\hat{\vec{L}}.\hat{\vec{S}}$, where g is the strength of the spin-orbit term. Note that there is no extra degree of freedom and the quantum numbers $\{n, l, m_l, m_s\}$ can still be used. But since components of $\hat{\vec{L}}$ and $\hat{\vec{S}}$ do not commute with $\hat{\vec{L}}.\hat{\vec{S}}$, they are no longer 'good'. We will see later that the 'good quantum numbers' can be chosen in this case as $\{n, l, j, m_j\}$, where j and m_j are the total angular momentum defined as $\hat{\vec{J}} = \hat{\vec{L}} + \hat{\vec{S}}$ and its z-projection, $\hat{J}_z$.

Orthogonality of eigen functions

We saw in Chap. 2, Sect. 2.3.1 that eigen vectors (hence eigen functions also) belonging to different eigen values of a Hermitian operator are orthogonal. Hence energy eigen functions (of the Hermitian Hamiltonian) belonging to different eigen values are orthogonal

$$\int_{-\infty}^{\infty} u_n^*(x) u_{n'}(x) \mathrm{d}x \propto \delta_{n,n'}$$

in one dimension (lower and upper limits of the integral will be a and b respectively if the system is restricted to the interval $[a, b]$). See below for three-dimensional problems. Usually each eigen function is normalized. The set of all eigen functions forms a complete set according to Chap. 2, Sect. 2.3.1. Hence this set is called the **complete orthonormal set (CONS)**.

Degeneracy

If an eigen value is the same for more than one combination of different quantum numbers, then the eigen value is called *degenerate*. Clearly, there cannot be any degeneracy in one dimension, as only one quantum number appears. Degenerate eigen functions belonging to different quantum numbers are linearly independent, but not automatically orthogonal. However they can be orthonormalized with respect to each quantum number. Thus for a three-dimensional spherically symmetric potential (see Chap 11, Sect. 11.2 and Chap 12, Sects. 12.1, 12.2, and 12.4) the energy eigen value depends on n and l, but not on m. The usual orthogonality relation is

$$\int u^*_{nlm}(\vec{r})u_{n'l'm'}(\vec{r})d^3r \propto \delta_{n,n'}\delta_{l,l'}\delta_{m,m'},$$

where the integration is over all space (unless restricted by applied conditions). However different linear combinations of $\{u_{nlm}(\vec{r}),\ m = -l, -l+1.\cdots, l\}$ corresponding to the same degenerate energy eigen value E_{nl} are in general not orthogonal.

Degeneracy arises due to symmetries of the Hamiltonian (which lead to separability of the Schrödinger equation in different coordinate systems). The $(2l+1)$-fold m-degeneracy (for a given l) for spherically symmetric Hamiltonian is called *rotational degeneracy*. This corresponds to an arbitrary choice of the polar axis of spherical polar coordinates for a spherically symmetric Hamiltonian. A particular linear combination of energy eigen functions with different m values corresponds to a particular orientation of the polar axis.

For the basic H-atom there is an additional l-degeneracy for a given n (with $l = 0, 1, \cdots, n-1$, for $n \geq 1$), which is related to the separability of the Schrödinger equation in parabolic coordinates (see Schiff 1968).

Choice of coordinate system according to symmetry of the problem

The choice of coordinate system in solving a particular Schrödinger equation is very important. The set of quantum numbers for one particular system will in general mix the corresponding set of quantum numbers for another system and an explicit relation between the two different sets may not be possible. For example, if the potential is a cubical box with rigid walls whose sides are aligned parallel to x-, y- and z-axes, then the Schrödinger equation separates immediately in these variables. We then have quantum numbers n_x, n_y and n_z, one for each of the three variables as in Chap. 9, Sect. 9.1. However if we choose the spherical polar coordinate system, it is not possible to separate the variables (r, θ, ϕ). Consequently, the (n, l, m) quantum numbers associated with the spherical polar system get mixed up. Similarly, if we consider the Schrödinger equation of a spherical hole with rigid walls (Chap. 12, Sect. 12.1), the equation in (r, θ, ϕ)-variables readily separates and corresponding quantum numbers can be easily defined. However in this case the Schrödinger equation in Cartesian variables cannot be separated.

Thus it is important to choose a coordinate system according to the *symmetry* of the problem. Also, separability of the Schrödinger equation in a particular coordinate system is related to degeneracy of the energy eigen values, as discussed above.

Parity of eigen functions

We can see that if a one-dimensional potential is symmetric about the origin, $V(-x) = V(x)$, then its eigen functions have a definite parity. Replacing x by $-x$ in the Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2u(x)}{\mathrm{d}x^2} + V(x)u(x) = Eu(x),$$

we have

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2u(-x)}{\mathrm{d}(-x)^2} + V(-x)u(-x) = Eu(-x),$$

which becomes using $V(-x) = V(x)$

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2u(-x)}{\mathrm{d}x^2} + V(x)u(-x) = Eu(-x).$$

Thus we see that both $u(x)$ and $u(-x)$ are eigen functions corresponding to the same eigen value E. Since in one dimension, there is no degeneracy, these must be proportional

$$u(-x) = Pu(x),$$

where P is a constant. Using the last relation twice

$$u(x) = u(-(-x)) = Pu(-x) = P^2u(x)$$
$$\text{Hence} \quad P^2 = 1, \quad P = \pm 1$$
$$\text{or} \quad u(-x) = \pm u(x)$$

Thus the eigen function has a definite parity if the potential is symmetric, $V(-x) = V(x)$. For non-symmetric potentials, the eigen functions have no definite parity.

In a three-dimensional case, the parity operation is $\vec{r} \to -\vec{r}$, which corresponds in polar coordinates to $(r, \theta, \phi) \to (r, \pi - \theta, \pi + \phi)$. Suppose the potential is even under the parity operation,

$$V(-\vec{r}\,) = V(\vec{r}\,).$$

Then the Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2 u(\vec{r}\,) + V(\vec{r}\,)u(\vec{r}\,) = Eu(\vec{r}\,)$$

under parity operation becomes

$$-\frac{\hbar^2}{2m}\nabla^2 u(-\vec{r}\,) + V(-\vec{r}\,)u(-\vec{r}\,) = Eu(-\vec{r}\,),$$

since ∇^2 remains invariant under parity. Using $V(-\vec{r}) = V(\vec{r})$. The last equation becomes

$$-\frac{\hbar^2}{2m}\nabla^2 u(-\vec{r}) + V(\vec{r})u(-\vec{r}) = Eu(-\vec{r}).$$

Thus both $u(-\vec{r})$ and $u(\vec{r})$ are eigen functions belonging to the eigen value E. But in three-dimension, degeneracy may occur depending on additional symmetry like spherical symmetry. When there is no symmetry other than the parity symmetry, we can still form two linear combinations

$$u_e(\vec{r}) \propto [u(\vec{r}) + u(-\vec{r})]$$
$$u_o(\vec{r}) \propto [u(\vec{r}) - u(-\vec{r})],$$

which are also eigen functions belonging to the same eigen value E, where $u_e(\vec{r})$ and $u_o(\vec{r})$ have even and odd parity respectively. Thus if the three-dimensional potential is invariant under parity, we can construct eigen functions which have definite parity. An example is provided by a particle confined within the space (with $V = 0$) limited by an ellipsoid of revolution, having rigid walls.

If a three-dimensional potential is spherically symmetric (which includes parity symmetry), the (θ, ϕ)-dependent part is a spherical harmonics $Y_{lm}(\theta, \phi)$, which has a parity $(-1)^l$ (see Chap. 10, Sect. 10.2). Hence the eigen function $u_{nlm}(\vec{r})$ has a definite parity $(-1)^l$ as in Eq. (11.20).

8.4 Free Particle Wave Function

Consider a *free particle* of mass m. Since the particle is free, i.e. it is not acted upon by any force, so we can take the potential $V(\vec{r}) = 0$ and from Eq. (4.4) we have (subscript f refers to a free particle)

$$-\frac{\hbar^2}{2m}\nabla^2\psi_f(\vec{r}, t) = i\hbar\frac{\partial\psi_f(\vec{r}, t)}{\partial t}. \tag{8.11}$$

Since the Hamiltonian is time-independent, time part separates

$$\psi_f(\vec{r}, t) = u_f(\vec{r})\exp(-\frac{i}{\hbar}Et) = u_f(\vec{r})\epsilon^{-i\omega t},$$

in which we write the separation constant E to be $E = \hbar\omega$. The spatial part $u_f(\vec{r})$ satisfies

$$-\frac{\hbar^2}{2m}\nabla^2 u_f(\vec{r}) = Eu_f(\vec{r}).$$

Its solution is easily found to be

$$u_f(\vec{r}) = \frac{1}{(2\pi)^{\frac{3}{2}}} e^{\frac{i}{\hbar}\vec{p}\cdot\vec{r}} = \frac{1}{(2\pi)^{\frac{3}{2}}} e^{i\vec{k}\cdot\vec{r}} \qquad \text{where } \vec{p} = \hbar\vec{k},$$

and $E = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m}$, $\vec{k}$ being the wave vector. We have adopted the δ-function normalization (see Sect. 3.5, Chap. 3). Hence the complete wave function of a free particle is

$$\psi_f(\vec{r}, t) = \frac{1}{(2\pi)^{\frac{3}{2}}} e^{i(\frac{\vec{p}}{\hbar}\cdot\vec{r} - \omega t)} = \frac{1}{(2\pi)^{\frac{3}{2}}} e^{i(\vec{k}\cdot\vec{r} - \omega t)}. \tag{8.12}$$

Corresponding free particle wave function in one dimension (x) is

$$\psi_f(x, t) == \frac{1}{(2\pi)^{\frac{1}{2}}} e^{i(kx - \omega t)}. \tag{8.13}$$

Applying the momentum operator $\hat{\vec{p}} = -i\hbar\vec{\nabla}$ on $\psi_f(\vec{r}, t)$ we see that it is an eigen function of $\hat{\vec{p}}$ corresponding to eigen value $\hbar\vec{k}$. Since the momentum is *exactly* specified, position-momentum uncertainty relation demands that position is completely unspecified. This is shown by the fact that $|\psi_f(\vec{r}, t)|^2$ is a constant, independent of $\vec{r}$. Thus the free particle wave function is a momentum eigen function. It is also called a traveling harmonic wave, since a particular phase travels in the direction $\vec{k}$ with the phase velocity $\frac{\omega}{k}$, without any change in the shape of the wave front. The wave front is an infinite plane perpendicular to the direction $\vec{k}$. Hence this wave is also called a plane wave.

8.5 Wave Packet and Its Motion

Equation (8.12) represents a free particle having a *definite* momentum $\vec{p} = \hbar\vec{k}$, satisfying the Schrödinger equation (8.11); hence $\frac{\hbar^2 k^2}{2m} = \hbar\omega$. However an actual experimental measurement yields a value within a range of experimental errors. For simplicity we consider motion in one dimension (say x), corresponding momentum being $\hbar k$. Suppose a momentum measurement yields the value $\hbar k$ in the range $\hbar(k_0 \pm \Delta k)$, with probability of finding momentum strongly peaking at $\hbar k_0$, within a small interval $\hbar(k_0 - \Delta k) \leq \hbar k \leq \hbar(k_0 + \Delta k)$. Since momentum is *not definite* now, the particle cannot be completely free and must satisfy one-dimensional form of the Schrödinger equation (4.4), which is linear and homogeneous in ψ. Hence a linear combination of its solutions is also a solution. We construct a superposition (called a wave packet) of wave functions (8.13) of various k values with a k-dependent amplitude $\phi(k)$

$$\psi(x,t) = \frac{1}{\sqrt{(2\pi)}} \int_{k_0-\Delta k}^{k_0+\Delta k} \phi(k) \mathrm{e}^{i(kx-\omega t)} dk. \tag{8.14}$$

Clearly the particle is not free anymore (hence we drop the suffix f) and is moving in a potential $V(x)$, which is very slowly varying with x (since the momentum is approximately known to be $\hbar k_0$ with a small uncertainty of Δk). This wave packet represents a particle at time t moving in x-direction with momentum lying between $\hbar(k_0 - \Delta k)$ and $\hbar(k_0 + \Delta k)$. Substituting

$$u = k - k_0,$$

the wave packet at time $t = 0$ is

$$\psi(x,0) = \frac{1}{\sqrt{(2\pi)}} \mathrm{e}^{ik_0 x} \int_{-\Delta k}^{\Delta k} \phi(u + k_0) \mathrm{e}^{iux} du. \tag{8.15}$$

Substitution of Eq. (8.14) in the one-dimensional form of Eq. (4.4), we notice that ω is now a function of k. Since $\phi(k)$ is strongly peaked about $k = k_0$, we expect that $\omega(k)$ will be a slowly varying function of k for values of k near k_0. Hence a Taylor series expansion about $k = k_0$ will be rapidly converging (see Merzbacher 1965). Retaining terms up to first order in Δk

$$\omega(k) = \omega(k_0) + \left.\frac{d\omega}{dk}\right|_{k=k_0} (k - k_0) + \cdots \approx \omega_0 + \omega_0' u,$$

where $\omega_0 = \omega(k_0)$ and $\omega_0' = \left.\frac{d\omega}{dk}\right|_{k=k_0}$. Substituting the approximate expression of $\omega(k)$ in Eq. (8.14) we have

$$\begin{aligned}
\psi(x,t) &= \frac{1}{\sqrt{2\pi}} \int_{-\Delta k}^{\Delta k} \phi(u + k_0) \mathrm{e}^{i[(u+k_0)x - (\omega_0 + \omega_0' u)t]} du \\
&= \mathrm{e}^{i(k_0 x - \omega_0 t)} \frac{1}{\sqrt{2\pi}} \int_{-\Delta k}^{\Delta k} \phi(u + k_0) \mathrm{e}^{i(x - \omega_0' t)u} du \\
&= \mathrm{e}^{i(k_0 x - \omega_0 t)} \mathrm{e}^{-ik_0(x - \omega_0' t)} \psi(x - \omega_0' t, 0) \\
&= \mathrm{e}^{-i(\omega_0 - \omega_0' k_0)t} \psi(x - \omega_0' t, 0),
\end{aligned} \tag{8.16}$$

in which we have used Eq. (8.15) with x replaced by $(x - \omega_0' t)$ in the last but one step. The final expression for $\psi(x,t)$ in Eq. (8.16) is seen to be the same as $\psi(x - \omega_0' t, 0)$, except for an unimportant phase factor. Thus the wave packet at time t is simply the

wave packet at time $t = 0$ which is shifted through a distance $x - \omega_0' t$. This shows that the wave packet moves with a speed ω_0' without much change in its shape (for possibility of change in shape see below). Then the group velocity of the wave packet is

$$v_g = \omega_0' = \left.\frac{d\omega}{dk}\right|_{k=k_0}.$$

If the particle is *nearly free*, then $\hbar\omega = \frac{\hbar^2 k^2}{2m} \approx \frac{\hbar^2 k_0^2}{2m}$ and v_g agrees with the velocity of the corresponding classical particle

$$v_g = \frac{\hbar k_0}{m} = \frac{p_c}{m} = v_c,$$

where p_c and v_c are respectively the momentum and velocity of the corresponding classical particle. This shows that the wave packet (with a fairly well defined momentum) follows the motion of the corresponding classical particle.

Note that Eq. (8.16) is valid if we stop the Taylor series expansion of $\omega(k)$ at the first order. Keeping higher orders will make a small change in shape of the wave packet as it moves with time. In general the wave packet spreads slowly as it progresses (Merzbacher, 1965).

As a particular example, let

$$\begin{aligned} \phi(k) &= 1 \qquad \text{for } k_0 - \Delta k \le k \le k_0 + \Delta k \\ &= 0 \qquad \text{otherwise,} \end{aligned}$$

so that probability of finding any value of momentum within this range is the same. Substituting this $\phi(k)$ in the second line of Eq. (8.16), and then using Eq. (8.15) and evaluating the integral, we have

$$\psi(x,t) = \sqrt{\frac{2}{\pi}} \frac{\sin\left((x - \omega_0' t)\Delta k\right)}{(x - \omega_0' t)} \, \mathrm{e}^{i(k_0 x - \omega_0 t)}. \tag{8.17}$$

This shows that the wave packet travels in $+x$-direction with a phase velocity of $\frac{\omega_0}{k_0}$. However it is not a constant amplitude wave. The amplitude has a damped sinusoidal nature (see Fig. 8.11), which is centered at $x_0 = \omega_0' t$ at time t (continuous curve), while at time $t = 0$ it is centered at the origin (dotted curve). The spatial extent of the packet can be estimated to be the width of the main peak at half its maximum. From Fig. 8.11 and Eq. (8.17) it is seen to be of the order of $\Delta x \sim \frac{\pi}{\Delta k}$. For values of x far from the center of this packet, the wave function is practically zero. Thus this is a traveling wave packet, whose spatial extension is $\sim \Delta x$. Since the uncertainty in momentum is $\Delta p = \hbar(2\Delta k)$, we have

$$\Delta x \Delta p \sim 2\pi\hbar = h,$$

in agreement with the uncertainty principle.

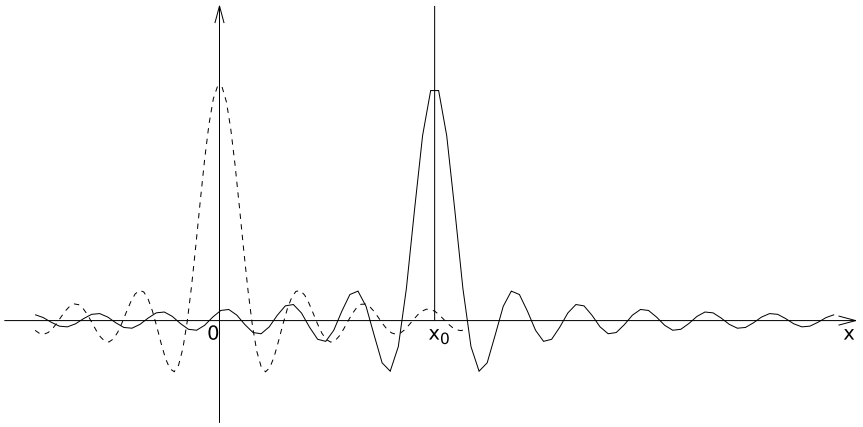


Fig. 8.11 Plot of the amplitude of the wave packet (8.17) for $t = \frac{x_0}{\omega_0'}$, which peaks at $x = x_0$ (continuous curve). The wave packet at $t = 0$ peaks at the origin (dotted curve). Thus the peak moves from $x = 0$ at time $t = 0$ to the position $x = x_0$ at time $t = \frac{x_0}{\omega_0'}$

In a similar fashion, a three-dimensional wave packet can be constructed as a linear combination of plane waves (8.12) with amplitude $\phi(\vec{k})$ as

$$\psi(\vec{r}, t) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int \phi(\vec{k}) e^{i(\vec{k}\cdot\vec{r} - \omega t)} d^3k. \tag{8.18}$$

As in one-dimensional case, if $\phi(\vec{k})$ is centered at $\vec{k} = \vec{k}_0$ and is negligible for $|k - k_0| \geq \Delta k$, so that its 'width' is $\sim 2\Delta k$, then we can see that $\psi(\vec{r}, t)$ also has a localized form, magnitude of whose extent is $\sim \frac{1}{\Delta k}$, in agreement with the uncertainty relation. Since Eq. (8.18) must satisfy the Schrödinger equation (4.4), we see that ω must be a function of k, which varies slowly with k at $k = k_0$. Expanding $\omega(k)$ about $k = k_0$ in a Taylor series we see (as in one-dimensional case) that the central peak of the wave packet travels with a phase velocity of magnitude $v_{ph} = \frac{\omega_0}{k_0}$ (where $\omega_0 = \omega(k_o)$) and the magnitude of group velocity of the wave packet is $v_{gr} = \frac{d\omega}{dk}$. Thus the center of the wave packet travels with a group velocity of magnitude $v_{gr,\, center} = \left.\frac{d\omega}{dk}\right|_{k=k_0} = \frac{\hbar k_0}{m}$, which is the classical speed of the particle with momentum $\hbar k_0$. Hence the wave packet representing a free particle with a fairly well specified momentum has a spatial extent and travels with a group velocity which is equal to the classical velocity of the particle, while the wave packet in general spreads gradually as it travels. Thus the quantum mechanical description agrees on the average with the classical description. A generalization of this result is called *Ehrenfest's theorem* and is presented in the following section.

General form of a wave packet moving in a potential

Suppose a particle of mass m moves in a time-independent potential $V(\vec{r})$ satisfying the Schrödinger equation

$$\Big[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})\Big]\psi(\vec{r},t) = i\hbar\frac{\partial\psi(\vec{r},t)}{\partial t}.$$

Then the wave function is separable in time and space as $\psi(\vec{r},t) = u(E,\vec{r})\,\mathrm{e}^{-\frac{iEt}{\hbar}}$ and $u(E,\vec{r})$ satisfies the time-independent Schrödinger equation

$$\Big[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})\Big]u(E,\vec{r}) = Eu(E,\vec{r}).$$

The separation constant E is the energy eigen value of the system. It is discrete for bound states and continuous for states that extend to infinity. We include E as an argument of $u(E,\vec{r})$, since the energy eigen function depends on E and forms a complete orthonormal set

$$\begin{aligned}
\text{Orthonormality:}\quad \int u^*(E,\vec{r})u(E',\vec{r})d^3r &= \delta(E-E'), \quad \text{for continuous E,}\\
&= \delta_{E,E'} \qquad \text{for discrete E}\\
\text{Completeness:}\quad \int u^*(E,\vec{r})u(E,\vec{r}\,')dE &= \delta^{(3)}(\vec{r}-\vec{r}\,') \quad \text{for continuous E,}\\
\sum_E u^*(E,\vec{r})u(E,\vec{r}\,') &= \delta^{(3)}(\vec{r}-\vec{r}\,') \quad \text{for discrete E.} \qquad (8.19)
\end{aligned}$$

Now consider a wave packet $\psi_{pac}(\vec{r},t_0)$ at time $t = t_0$. According to the expansion postulate, this wave packet can be expanded in the complete orthonormal set of energy eigen functions

$$\begin{aligned}
\psi_{pac}(\vec{r},t_0) &= \int \phi(E)\,u(E,\vec{r})\,\mathrm{e}^{-i\frac{E}{\hbar}t_0}dE \quad \text{for continuous E,}\\
&= \sum_E \phi(E)\,u(E,\vec{r})\,\mathrm{e}^{-i\frac{E}{\hbar}t_0} \quad \text{for discrete E,} \qquad (8.20)
\end{aligned}$$

where $\phi(E)$ is the expansion coefficient. Using orthonormality, Eq. (8.19) we have for $\phi(E)$

$$\phi(E)\,\mathrm{e}^{-i\frac{E}{\hbar}t_0} = \int \psi_{pac}(\vec{r},t_0)\,u^*(E,\vec{r})d^3r.$$

Substituting for $\phi(E)$ from above in Eq. (8.20) with t_0 replaced by an arbitrary time t (and variables of integration E and $\vec{r}$ replaced by E' and $\vec{r}\,'$ respectively), we get

$$\psi_{pac}(\vec{r},t) = \int_{E'} \Big[\int_{\vec{r}\,'} \psi_{pac}(\vec{r}\,',t_0)\, u^*(E',\vec{r}\,')d^3r' \Big] u(E',\vec{r}\,)\, \mathrm{e}^{-i\frac{E'}{\hbar}(t-t_0)} dE'$$

for continuous E,

$$= \sum_{E'} \Big[\int_{\vec{r}\,'} \psi_{pac}(\vec{r}\,',t_0)\, u^*(E',\vec{r}\,')d^3r' \Big] u(E',\vec{r}\,)\, \mathrm{e}^{-i\frac{E'}{\hbar}(t-t_0)}$$

for discrete E. (8.21)

This gives the propagation of the wave packet with time. In general the wave packet will spread as it travels. We will see a specific case about the motion of a wave packet in a harmonic oscillator well in Chap. 9, Sect. 9.5. In that case the wave packet does not spread with time. This is because the initial wave packet (chosen in the case) is the ground state in the well, which is the minimum uncertainty packet.

8.6 Ehrenfest's Theorem

Although individual eigen values and eigen states usually do not follow laws of classical physics, the overall motion of wave packets obeys classical laws. This is commonly known as *Ehrenfest's theorem*, which states that the expectation values of dynamical variables for a wave packet obey laws of classical physics. We will demonstrate it by examples.

For simplicity, first consider a one-dimensional motion of a wave packet $\psi(x,t)$ satisfying the Schrödinger equation

$$\Big[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \Big] \psi(x,t) = i\hbar \frac{\partial \psi(x,t)}{\partial t}, \tag{8.22}$$

where ψ is assumed normalized. The time derivative of $\langle x \rangle$ is

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\langle x \rangle &= \frac{\mathrm{d}}{\mathrm{d}t} \int_{-\infty}^{\infty} \psi^*(x,t)\, x\, \psi(x,t)\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \psi^*(x,t)\, x \Big(\frac{\partial \psi(x,t)}{\partial t} \Big) \mathrm{d}x + \int_{-\infty}^{\infty} \Big(\frac{\partial \psi^*(x,t)}{\partial t} \Big) x\, \psi(x,t)\mathrm{d}x \\
&= \frac{i\hbar}{2m} \Big[\int_{-\infty}^{\infty} \psi^*(x,t)\, x \frac{\partial^2 \psi(x,t)}{\partial x^2} \mathrm{d}x - \int_{-\infty}^{\infty} \frac{\partial^2 \psi^*(x,t)}{\partial x^2}\, x\, \psi(x,t)\mathrm{d}x \Big] \\
&= \frac{i\hbar}{2m} \Big[1^{st} \text{ term } - \text{2nd term } \Big],
\end{aligned} \tag{8.23}$$

using Eq. (8.22). Evaluating the 2nd term of Eq. (8.23) by partial integrations in first and third steps of the following (and suppressing arguments of ψ) we have

$$\begin{aligned}
\text{2nd term} &= \Big(\frac{\partial\psi^*}{\partial x}\Big)x\psi\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\frac{\partial}{\partial x}\Big(x\psi\Big)\frac{\partial\psi^*}{\partial x}\mathrm{d}x \\
&= 0 - \int_{-\infty}^{\infty}\Big(\psi + x\frac{\partial\psi}{\partial x}\Big)\frac{\partial\psi^*}{\partial x}\mathrm{d}x \qquad \text{(since } \psi \text{ vanishes at } \pm\infty\text{)} \\
&= -\Big(\psi + x\frac{\partial\psi}{\partial x}\Big)\psi^*\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty}\Big(\frac{\partial\psi}{\partial x} + \frac{\partial\psi}{\partial x} + x\frac{\partial^2\psi}{\partial x^2}\Big)\psi^*\mathrm{d}x \\
&= 0 + 2\int_{-\infty}^{\infty}\psi^*\frac{\partial\psi}{\partial x}\mathrm{d}x + \int_{-\infty}^{\infty}\psi^* x\frac{\partial^2\psi}{\partial x^2}\mathrm{d}x.
\end{aligned}$$

Substituting this result back in Eq. (8.23), we have

$$\frac{\mathrm{d}\langle x\rangle}{\mathrm{d}t} = \frac{1}{m}\int_{-\infty}^{\infty}\psi^*\Big(-i\hbar\frac{\partial}{\partial x}\Big)\psi\mathrm{d}x = \frac{\langle p_x\rangle}{m}. \tag{8.24}$$

This is similar to the classical relation for velocity as momentum divided by mass

$$\frac{\mathrm{d}x_{cl}}{\mathrm{d}t} = \frac{p_{x,cl}}{m},$$

where the subscript cl stands for the corresponding classical quantity.

As another example, we investigate the equivalence of the classical force

$$\vec{F}_{cl} = \frac{\mathrm{d}\vec{p}_{cl}}{\mathrm{d}t} = -\vec{\nabla}V_{cl}(\vec{r}\,). \tag{8.25}$$

Again for simplicity, we consider only one-dimensional motion. We have for a wave packet $\psi(x,t) \equiv \psi$

$$\begin{aligned}
\frac{\mathrm{d}\langle p_x\rangle}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}\psi^*\Big(-i\hbar\frac{\partial}{\partial x}\Big)\psi\mathrm{d}x \\
&= \int_{-\infty}^{\infty}\psi^*\Big(-i\hbar\frac{\partial}{\partial x}\Big)\frac{\partial\psi}{\partial t}\mathrm{d}x + \int_{-\infty}^{\infty}\frac{\partial\psi^*}{\partial t}\Big(-i\hbar\frac{\partial\psi}{\partial x}\Big)\mathrm{d}x.
\end{aligned}$$

Substituting for the time derivatives from Eq. (8.22), we have

$$\frac{\mathrm{d}\langle p_x\rangle}{\mathrm{d}t} = \int_{-\infty}^{\infty} \psi^* \frac{\partial}{\partial x}\Big[\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} - V\psi\Big]\mathrm{d}x + \int_{-\infty}^{\infty}\Big[-\frac{\hbar^2}{2m}\frac{\partial^2\psi^*}{\partial x^2} + V\psi^*\Big]\frac{\partial\psi}{\partial x}\mathrm{d}x.$$

As before, the integral involving $\frac{\partial^2\psi}{\partial x^2}$ of the first integral can be evaluated, integrating twice by parts. It then cancels with the similar term of the second integral. Hence

$$\frac{\mathrm{d}\langle p_x\rangle}{\mathrm{d}t} = \int_{-\infty}^{\infty} \psi^*\Big[\frac{\partial}{\partial x}\Big(-V\psi\Big) + V\frac{\partial\psi}{\partial x}\Big]\mathrm{d}x = \int_{-\infty}^{\infty}\psi^*\Big(-\frac{\partial V}{\partial x}\Big)\psi\,\mathrm{d}x = \Big\langle -\frac{\partial V}{\partial x}\Big\rangle. \tag{8.26}$$

This relation is similar to the x-component of the classical force relation eq (8.25). We have obtained both Eqs. (8.24) and (8.26) in x-direction only. We can follow the steps for y-direction and z-directions also. Combining them we get the vector relations. Alternatively, we can use differential vector identities (although more complicated) to arrive at the same results.

Thus it is seen that if the wave packet has a finite extent, the overall motion of the packet is governed by the classical laws. This provides the transition from quantum mechanical results to the classical results, when only average values over wave packets are needed and microscopic details are not sought (this is called *the correspondence limit*). It is the basis of the *correspondence principle*, which states that quantum mechanical results go over to the corresponding classical results, whenever the situation is such that the size and structure of the wave packet can be disregarded for the classical variable.

Expectation value of momentum vanishes for a bound state

As a consequence of Ehrenfest's theorem, we can see that the expectation value of momentum vanishes for a bound state. Three-dimensional generalization of Eq. (8.24) gives

$$\langle\,\vec{p}\,\rangle = m\frac{\mathrm{d}\langle\,\vec{r}\,\rangle}{\mathrm{d}t}. \tag{8.27}$$

Now

$$\langle\,\vec{r}\,\rangle = \int \psi_E^*(\vec{r},t)\,\vec{r}\;\psi_E(\vec{r},t)\,d^3r.$$

For a bound state corresponding to energy eigen value E (which must be real)

$$\psi_E(\vec{r},t) = u_E(\vec{r})\;\mathrm{e}^{-i\frac{E}{\hbar}t}. \tag{8.28}$$

Substituting in the expression for $\langle \vec{r} \rangle$, we have

$$\langle \vec{r} \rangle = m \int u_E^*(\vec{r})\, \vec{r}\, u_E(\vec{r})\, d^3r \quad \rightarrow \text{independent of } t.$$

Hence from Eq. (8.27), we have $\langle \vec{p} \rangle = 0$. This is true for a bound state and can be understood physically, since a bound state is a *localized state* and cannot be moving as a whole. However it has time dependence according to Eq. (8.28) as a 'standing wave'. Hence a bound state is also called a *stationary state*.

8.7 Problems

1. For the general potential shown in Fig. 8.2, provide arguments to show that there will be *at least one* bound state in this one-dimensional potential, independent of the depth or width of the potential.
2. Consider an arbitrary one-dimensional potential $V(x)$, which has a minimum $V_{\min}$ at a finite value of x and is finite (not zero) both for $x \rightarrow +\infty$ and $x \rightarrow -\infty$. Show that there will be one or more bound states having energy E lying in the range $V_{\min} < E < V_M$, where V_M is the smaller of $V(-\infty)$ and $V(+\infty)$. Convince yourself that there will be *at least one bound state* in this potential. Show also that there will be continuous energy eigen values for $E > V_M$.
3. Consider a finite square well potential (Chap. 9, Sect. 9.2). Estimate the number of bound states in this well according to Eq. (8.10) and compare with the actual number for several values of $\frac{2mV_0a^2}{\hbar^2}$ from Figs. 9.3 and 9.4, together with Eq. (9.16).
4. A particle of mass m moves in a potential $V(\vec{r})$. Use Heisenberg equation of motion to calculate

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\vec{r}}.\hat{\vec{p}}.$$

Then obtain the expectation value of this quantity for the normalized *stationary* state $\psi(\vec{r}, t)$ and show that

$$\langle \hat{T} \rangle = \frac{1}{2} \langle \hat{\vec{r}}.\hat{\vec{\nabla}} V \rangle,$$

where $\hat{T}$ is the kinetic energy operator. This is known as Virial theorem.

References

Merzbacher, E.: Quantum Mechanics. John Wiley & Sons Inc., New York (1965)
Schiff, L.I.: Quantum Mechanics, 3rd edn. McGraw-Hill Book Company Inc., Singapore (1968)

Chapter 9
One-Dimensional Potentials



Abstract In this chapter general procedure for solving simple one-dimensional problems in position coordinate space is discussed. As typical examples we discuss: infinite and finite square well, harmonic oscillator well, infinite well with a delta function, quasi-bound state in a delta function barrier. Motion of a wave packet in a harmonic oscillator well is also discussed.

Keywords Exact soluions of Schrödinger equation in one dimension · Infinite and finite square wells · General procedure · Harmonic oscillator · Wave packet in HO · Delta function potential

In this chapter we will discuss the solution of Schrödinger equation for simple one-dimensional potentials. These potentials are important for simplified and mathematically idealized physical situations. The resulting differential equations are standard ones with known solutions. For the simple harmonic oscillator, instead of referring to the standard Hermite differential equation, we will use the Frobenius method, to demonstrate how the boundary conditions lead to discrete energy eigen values, as discussed in Chap. 8.

9.1 A Particle in a Rigid Box

As a simple example, we consider a particle of mass m in a rigid one-dimensional box of width $2a$. By *rigid box* we mean the box has rigid walls at a separation of $2a$, so that the particle is confined within the box and cannot cross the walls, no matter how energetic it is. Thus the potential at the walls becomes suddenly infinite. We take the potential inside the box to be zero and the walls placed symmetrically about the origin (as in Schiff 1968), so that

$$\begin{aligned} V(x) &= 0 \quad \text{for } -a \le x \le a \quad \text{(Region I)} \\ &= \infty \quad \text{for } |x| > a \quad \text{(Region II)} \end{aligned}$$

The potential is shown in Fig. 9.1. The Schrödinger equation is

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5_9

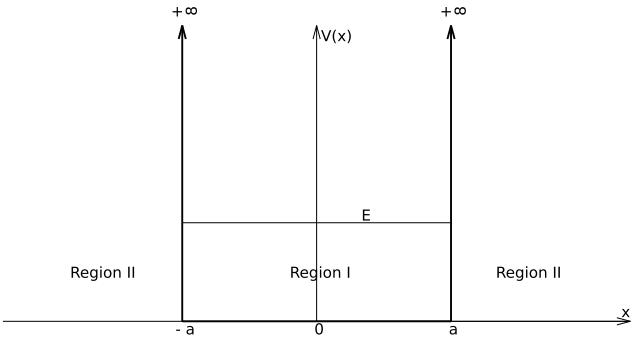


Fig. 9.1 Plot of the potential of a rigid box

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2u(x)}{\mathrm{d}x^2} + V(x)u(x) = Eu(x).$$

In Region I, $V(x) = 0$, hence

$$\frac{\mathrm{d}^2u(x)}{\mathrm{d}x^2} + \alpha^2 u(x) = 0, \quad \text{where } \alpha = +\sqrt{\frac{2mE}{\hbar^2}}.$$

This is a standard equation. The general solution is

$$u(x) = A\sin\alpha x + B\cos\alpha x$$

Boundary conditions at $x = \pm a$ require that the wave function vanish at $x = \pm a$

$$u(+a) = A\sin\alpha a + B\cos\alpha a = 0, \tag{9.1}$$

and

$$u(-a) = -A\sin\alpha a + B\cos\alpha a = 0, \tag{9.2}$$

From Eqs. (9.1) and (9.2), we have

$$A\sin\alpha a = 0, \tag{9.3}$$

$$B\cos\alpha a = 0. \tag{9.4}$$

Now $\sin\alpha a$ and $\cos\alpha a$ cannot both be zero. Again both A and B cannot vanish (since then $u(x) = 0$, which is a trivial solution), hence we have two possibilities, giving rise to two distinct classes of solutions:
Class I
$A = 0$, and $\cos\alpha a = 0$ hence $\alpha a = \frac{n\pi}{2}$ with $n =$ odd integer.
Class II
$B = 0$, and $\sin\alpha a = 0$ hence $\alpha a = \frac{n\pi}{2}$ with $n =$ even integer, but $n \neq 0$.
The restriction $n \neq 0$ is needed, since otherwise $u(x) = 0$ identically. In either case, we have

$\alpha a = \frac{n\pi}{2}$, with $n =$ an integer, odd or even, but $n \neq 0$. Thus

$$\alpha = \frac{n\pi}{2a} = \sqrt{\frac{2mE}{\hbar^2}} \;\Rightarrow\; E_n = \frac{n^2\pi^2\hbar^2}{8ma^2}, \quad n = 1, 2, 3, \cdots . \tag{9.5}$$

Thus we get an infinite set of discrete energy eigen values, characterized by the *quantum number* n. E_n increases with n and the minimum $E_1 = \frac{\pi^2\hbar^2}{8ma^2}$ is for $n = 1$ and is the *ground state*. The eigen function corresponding to the eigen value n is

$$\text{Class I}: \quad u_n(x) = B\cos\left(\frac{n\pi x}{2a}\right), \quad n = 1, 3, 5, \cdots .$$
$$\text{Class II}: \quad u_n(x) = A\sin\left(\frac{n\pi x}{2a}\right), \quad n = 2, 4, 6, \cdots .$$

The normalization constant (A or B) can be found from the normalization condition

$$\int_{-a}^{a} |u_n(x)|^2 \mathrm{d}x = 1.$$

For Class I and Class II solutions, these integrals are

$$\text{Class I}: \quad |B|^2 \int_{-a}^{a} \left|\cos\left(\frac{n\pi x}{2a}\right)\right|^2 \mathrm{d}x = |B|^2 a,$$
$$\text{Class II}: \quad |A|^2 \int_{-a}^{a} \left|\sin\left(\frac{n\pi x}{2a}\right)\right|^2 \mathrm{d}x = |A|^2 a.$$

Hence

$$A = B = \frac{1}{\sqrt{a}}.$$

Thus the normalized eigen functions are

$$\text{Class I}: \quad u_n(x) = \frac{1}{\sqrt{a}}\cos\left(\frac{n\pi x}{2a}\right), \quad n = 1, 3, 5, \cdots . \tag{9.6}$$
$$\text{Class II}: \quad u_n(x) = \frac{1}{\sqrt{a}}\sin\left(\frac{n\pi x}{2a}\right), \quad n = 2, 4, 6, \cdots . \tag{9.7}$$

Note:

1. We get an infinite set of discrete energy eigen values with quantum number $n = 1, 2, \cdots, \infty$. The ground state ($n = 1$) has no nodes, the n-th *excited* state [having quantum number $(n + 1)$] has n nodes, in agreement with our general discussion in Chap. 8, Sect. 8.2.

2. The energy eigen values are consistent with the uncertainty relation Eq. (5.12) as seen below. Since the particle can be anywhere in $-a < x < a$, we set $\Delta a \approx a$ (in place of a specific definition for the variance)

$$\Delta x \Delta p_x \geq \frac{\hbar}{2} \quad \Rightarrow \quad \Delta p_x \geq \frac{\hbar}{2a}.$$

Now, since for a bound state $\langle p_x \rangle = 0$ (see Chap. 8, Sect. 8.6)

$$(\Delta p_x)^2 \equiv \langle p_x^2 \rangle - \langle p_x \rangle^2 = \langle p_x^2 \rangle. \qquad \text{Hence } \langle p_x^2 \rangle \geq \frac{\hbar^2}{4a^2}.$$

Then

$$E = \frac{\langle p_x^2 \rangle}{2m} \geq \frac{\hbar^2}{8ma^2}.$$

This is consistent with Eq. (9.5).

3. $u_n(x)$ is even if n is odd and odd if n is even. Thus

$$u_n(-x) = (-1)^{(n-1)}\, u_n(x).$$

It is said that Class I eigen functions have *even parity* and Class II eigen functions have *odd parity*. Since the potential is symmetric about the origin, we must have the probabilities of finding the particle at x and at $-x$ the same, i.e. $|u_n(-x)|^2 = |u_n(x)|^2$. This is true if the eigen function has a definite parity. These are consistent with Chap. 8, Sect. 8.3.

4. Instead of choosing our origin at the middle of the well, we could choose it at one end, say the left end, so that

$$\begin{aligned} V(x) &= +\infty \quad \text{for } x < 0 \\ &= 0 \quad \text{for } 0 \leq x \leq 2a \\ &= +\infty \quad \text{for } x > 2a. \end{aligned}$$

In this case the potential is no more symmetric, but the physical case remains the same. It can be seen (see Problems at the end of this chapter) that the eigen values remain the same, but the eigen functions are shifted along x by $-a$, as expected. There is no sense in saying that the wave functions have any symmetry about the origin, since in this choice, the particle cannot be in the region $x < 0$. However we can see that the eigen functions are either symmetric or anti-symmetric *about the middle of the well,* $x = a$.

If a potential is symmetric, the effort in solving the Schrödinger equation can be reduced by using the symmetry property.

9.2 A Particle in a Finite Square Well

Next we consider a one-dimensional box with permeable walls of height V_0. A classical particle with energy $E < V_0$ will be confined within the box. But a quantum mechanical particle has a finite probability of crossing the walls. The potential is shown in Fig. 9.2 (which also indicates the Regions):

$$\begin{aligned} V(x) &= 0 \quad \text{for} \quad -a < x < a \\ &= V_0 \quad \text{for} \quad |x| \geq a. \end{aligned}$$

We have chosen the origin such that the potential is symmetric (Schiff 1968) and use this symmetry to reduce the effort needed (see Problems at the end of this chapter). Since the potential is symmetric about the origin, the eigen functions will have a definite parity, either even or odd. For a bound state $E < V_0$. Using the symmetry, the Schrödinger equation is solved only (say) for $x \geq 0$. In Region I,

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2 u_{_I}(x)}{\mathrm{d}x^2} = E u_{_I}(x).$$

As before, the general solution is

$$u_{_I}(x) = A \sin \alpha x + B \cos \alpha x \quad \text{where } \alpha = \sqrt{\frac{2mE}{\hbar^2}}. \tag{9.8}$$

In Region II,

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2 u_{_{II}}(x)}{\mathrm{d}x^2} + V_0 u_{_{II}}(x) = E u_{_{II}}(x).$$

or

$$\frac{\mathrm{d}^2 u_{_{II}}(x)}{\mathrm{d}x^2} - \beta^2 u_{_{II}}(x) = 0,$$

where

$$\beta = \sqrt{\frac{2m(V_0 - E)}{\hbar^2}}.$$

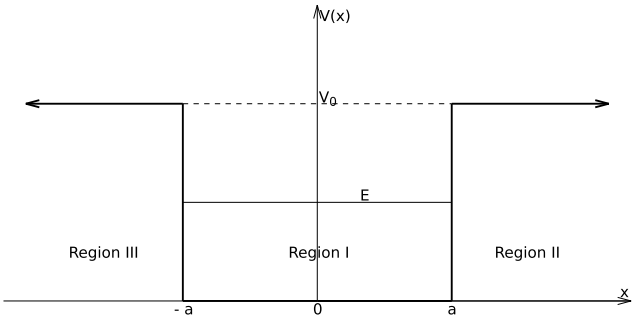


Fig. 9.2 Plot of the finite square well potential

The general solution is a linear combination of an outwardly growing ($\propto e^{\beta x}$) and a decaying ($\propto e^{-\beta x}$) exponential functions. Since the wave function must be finite at large x, the constant in front of the growing exponential must vanish. Hence an acceptable general solution in Region II is

$$u_{II}(x) = Ce^{-\beta x}, \tag{9.9}$$

We now separate even and odd class of eigen functions.

Class I: Even solutions
We see from Eq. (9.8) that for the even solutions, we must set $A = 0$. Then $u_I(x) = B\cos\alpha x$. Next, we have to match $u(x)$ and $\frac{du(x)}{dx}$ at $x = a$ only. The arbitrary constants cancel out if we match $\frac{1}{u}\frac{du}{dx}$ (called *continuity of log derivative*) and get

$$\alpha \tan \alpha a = \beta. \tag{9.10}$$

Since α and β contain the unknown E, this ia a transcendental equation for the eigen value E for even eigen functions. The arbitrary constant C can be related to B by continuity of the eigen function at $x = a$

$$B\cos\alpha a = Ce^{-\beta a}, \tag{9.11}$$

and finally the last unknown constant B can be found from the normalization condition

$$\int_{-\infty}^{+\infty} |u(x)|^2 dx = 2\left[\int_0^a |u_I(x)|^2 dx + \int_a^{+\infty} |u_{II}(x)|^2 dx\right] = 1 \tag{9.12}$$

$$= 2\left[|B|^2 \int_0^a \cos^2\alpha x \, dx + |B|^2 \cos^2\alpha a \, e^{2\beta a} \int_a^{\infty} e^{-2\beta x} \, dx\right]$$

$$= |B|^2\left[a + \tfrac{1}{\beta}\right].$$

To get the last line, Eq. (9.10) has been used. Hence

$$B = \left(a + \frac{1}{\beta}\right)^{-\frac{1}{2}}, \quad C = \left(a + \frac{1}{\beta}\right)^{-\frac{1}{2}} \frac{\alpha}{\sqrt{\alpha^2 + \beta^2}} e^{\beta a}.$$

and the normalized even eigen functions are given by Eqs. (9.8) and (9.9) as [for the even class, we have $u(-x) = u(x)$]

$$\begin{aligned} u(x) &= Ce^{\beta x} \quad \text{for } x < -a \\ &= B\cos\alpha x \quad \text{for } -a < x < a \\ &= Ce^{-\beta x} \quad \text{for } x > a, \end{aligned}$$

where B and C are given as above.

Class II: Odd solutions
For this $B = 0$ and $u_I(x) = A \sin \alpha x$. As before we match $\frac{1}{u}\frac{\mathrm{d}u}{\mathrm{d}x}$ and get

$$\alpha \cot \alpha a = -\beta. \tag{9.13}$$

This ia again a transcendental equation for the eigen value E for odd eigen functions. The arbitrary constant C can be related to A by continuity of the eigen function at $x = a$

$$A \sin \alpha a = C\mathrm{e}^{-\beta a}, \tag{9.14}$$

and finally as in the previous case, the last unknown constant A can be found from the normalization condition

$$\int_{-\infty}^{+\infty} |u(x)|^2 \mathrm{d}x = 2\left[\int_0^a |u_I(x)|^2 \mathrm{d}x + \int_a^{+\infty} |u_{II}(x)|^2 \mathrm{d}x\right] = 1. \tag{9.15}$$

Expressing C in terms of A using Eq. (9.14), the normalization constant A is determined as for the even class. In this case the eigen energy equation (9.13) for the odd class is to be used. The final result is again

$$A = \left(a + \frac{1}{\beta}\right)^{-\frac{1}{2}}, \quad C = \left(a + \frac{1}{\beta}\right)^{-\frac{1}{2}} \frac{\alpha}{\sqrt{\alpha^2 + \beta^2}}\, \mathrm{e}^{\beta a}.$$

Substituting these the normalized odd eigen functions are given by

$$\begin{aligned} u(x) &= -C\mathrm{e}^{\beta x} \quad \text{for } x < -a \\ &= A \sin \alpha x \quad \text{for } -a < x < a \\ &= C\mathrm{e}^{-\beta x} \quad \text{for } x > a, \end{aligned}$$

with A and C as given above.

Energy eigen values of even parity states
Energy equations are transcendental equations and cannot be solved analytically. They must be solved numerically or graphically. We discuss the graphical method (Schiff 1968) to get eigen values.
Let

$$\xi = \alpha a, \quad \eta = \beta a. \quad \text{Then } \xi^2 + \eta^2 = \frac{2mV_0a^2}{\hbar^2} \equiv R_\odot^2 = \text{constant} \tag{9.16}$$

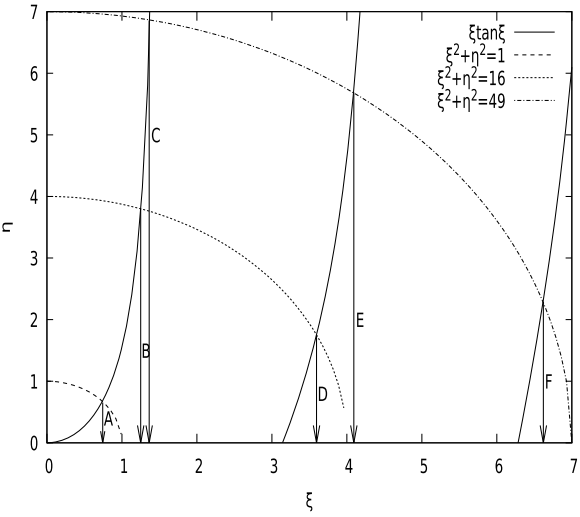


Fig. 9.3 Plot of $\xi \tan \xi$ and $\xi^2 + \eta^2 =$ constant in a ξ-η plot for even parity states. The ξ-values at the head of arrows A, B, C, D, E, F give respectively the eigen energies $E_1^{(1)}$, $E_1^{(2)}$, $E_1^{(3)}$, $E_3^{(2)}$, $E_3^{(3)}$ and $E_5^{(3)}$, according to Eqs. (9.16) and (9.8) and the identification $E_n^{(i)}$ given after Eq. (9.17)

represents a circle of radius $R_\odot = \sqrt{\frac{2mV_0a^2}{\hbar^2}}$ in the first quadrant in (ξ, η)-plane (we need only positive values of ξ and η). For the even class, the energy equation, Eq. (9.10), becomes in terms of these variables

$$\xi \tan \xi = \eta. \tag{9.17}$$

Similarly for odd parity states we have from Eq. (9.13)

$$\xi \cot \xi = -\eta.$$

Figure 9.3 shows plots of Eqs. (9.17) and (9.16). We consider three typical cases with $\xi^2 + \eta^2 = 1, 16$ and 49, corresponding to potential parameters $V_0a^2 = \frac{\hbar^2}{2m}$, $\frac{16\hbar^2}{2m}$ and $\frac{49\hbar^2}{2m}$ respectively. Corresponding energy eigen values of successive even parity states are denoted by $E_n^{(1)}$, $E_n^{(2)}$ and $E_n^{(3)}$ respectively, with $n = 1, 3, 5, \cdots$. We label the successive eigen values (for the i-th potential parameter) as $E_n^{(i)}$, with $n = 1, 3, 5, \cdots$ for the even class and $n = 2, 4, 6, \cdots$ for the odd class solutions. The ξ-value of the point of intersection of these curves is a solution of Eq. (9.17) [equivalent to Eq. (9.10)] and Eq. (9.16). Plot of Eq. (9.17) has an infinite number of branches with the first branch starting from the origin. From the figure, we can see that there will be one solution *even for very small values of* $\frac{2mV_0a^2}{\hbar^2}$. Thus at least one solution exists for all values of parameters of the potential and mass of the particle. This is true in one-dimension, as noted in Chap. 8, Sect. 8.2. As V_0a^2 increases for a given m, the radius of the circle increases and crosses *two* branches of Eq. (9.17) for $V_0a^2 > \frac{\pi^2\hbar^2}{2m}$(corresponding to radius $R_\odot > \pi \approx 3.142$), giving rise to *two even class solutions*. The new one is an excited state with even parity. For $V_0a^2 > \frac{4\pi^2\hbar^2}{2m}$

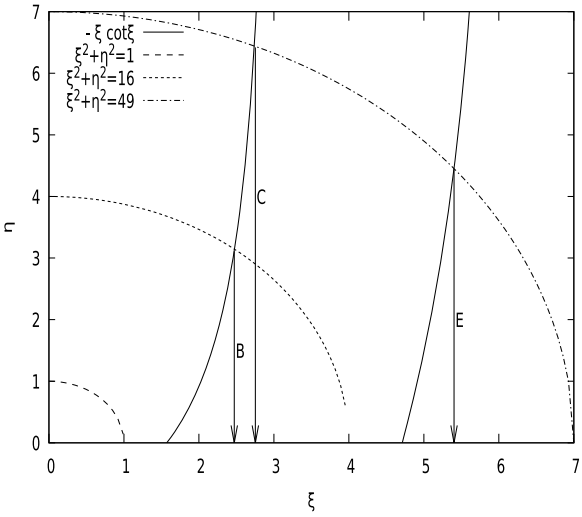


Fig. 9.4 Plot of $-\xi \cot \xi$ and $\xi^2 + \eta^2 = \text{constant}$ in a ξ-η plot for odd parity states. The ξ-values at the head of arrows B, C and E give respectively the eigen energies $E_2^{(2)}$, $E_2^{(3)}$ and $E_4^{(3)}$, according to Eqs. (9.16) and (9.13). Note that there is no solution for $\xi^2 + \eta^2 < \frac{\pi^2}{4}$

(corresponding to radius $R_\odot > 2\pi \approx 6.285$), there will be *three* even class solutions, and so on. Thus the number of bound states will increase as $V_0 a^2$ increases.

Energy eigen values of odd parity states

These are given by Eqs. (9.13) and (9.16). Proceeding in the same way as above we again get a number of energy eigen values for a given value of $V_0 a^2$. As before, energy eigen values are given by the intersection of the curves $-\xi \cot \xi$ and $\xi^2 + \eta^2 =$ constant. However in this case, the first branch of $-\xi \cot \xi$ (in the first quadrant, since both ξ and η should be positive) begins at $R_\odot = \frac{\pi}{2}$. Hence for $V_0 a^2 < \frac{\pi^2 \hbar^2}{8m}$ there is no odd parity solution (see Fig. 9.4). Note that solutions with even parity and odd parity appear alternately for a given choice of potential parameters and mass.

Estimation of the number of bound states in a finite square well potential

From Chap. 8, Sect. 8.3, the number of bound states in a potential $V(x)$ can be estimated as

$$n_{\max} = \text{integer part of} \left[\int_{\text{interval of } x} \sqrt{\frac{2m}{\pi^2 \hbar^2} (E_{n_{\max}} - V(x))}\ \mathrm{d}x \right] + 1,$$

where $E_{n_{\max}}$ is the energy of the uppermost bound state. For the potential of Fig. 9.2, we have $E_{n_{\max}}$ slightly less than V_0 and $V(x) = 0$ in the interval $-a < x < a$. Taking $E_{n_{\max}} \approx V_0$

$$n_{\max} = \text{integer part of } \left[\int_{-a}^{a} \sqrt{\frac{2m}{\pi^2\hbar^2} V_0}\ \mathrm{d}x \right] + 1$$

$$= \text{integer part of } \left[\sqrt{\frac{8m}{\pi^2\hbar^2} V_0 a^2} \right] + 1.$$

The total number of bound states is obtained by adding the number of even states (intersections of Fig. 9.3) and number of odd states (intersections of Fig. 9.4), for a given value of $R_\odot^2 = \frac{2mV_0a^2}{\hbar^2}$ according to Eq. (9.16). From these two figures it is found that the estimate is correct for all three cases chosen, *viz.*, $R_\odot^2 = 1$, 16 and 49. The estimate becomes better as $R_\odot^2$ increases. Clearly, $n_{\max} \to \infty$ as $V_0 \to \infty$, as for the infinite square well.

9.3 General Procedure for Bound States

So far we considered potentials with rectangular wells, for which the solution of the differential equation is well known. For more general potentials we can follow a general procedure. For bound state motions in more than one dimension, an attempt should be made to choose a coordinate system in which the Schrödinger equation separates into single variable equations. Then one of the second order differential equations in a particular variable contains the energy eigen value. This equation is to be solved according to the following procedure.

1. Introducing a suitable dimensionless variable (say ξ) express the Schrödinger equation in a dimensionless form. The dimensionless constant involving E is denoted by λ.
2. If large values of ξ are included in the physical interval, obtain the asymptotic (large $|\xi|$) form of the equation. Let its solution, which vanishes as $|\xi| \to \infty$, be $\phi_\infty(\xi)$. Express the wave function, $u(\xi)$, as a product of $\phi_\infty(\xi)$ times a function still undetermined [say $v(\xi)$], so that $u(\xi) = \phi_\infty(\xi)v(\xi)$.
3. If the origin is included in the domain for ξ and it is a singular point, then extract the $\xi \to 0$ behavior of the effective Schrödinger equation in ξ. Choose its solution which is acceptable (leading to finite $u(x)$ as $x \to 0$). Let it be $\phi_0(\xi)$. In this case, express $u(\xi) = \phi_0(\xi)\phi_\infty(\xi)v(\xi)$.
4. Substitute this form in the effective Schrödinger equation and extract the second order differential equation for $v(\xi)$.
5. If this equation is identified to be a standard differential equation (SDE) with known solutions, the wave function is known in terms of the standard solution multiplied by $\phi_\infty(\xi)$ or by $\phi_0(\xi)\phi_\infty(\xi)$, as the case may be. From the condition that this wave function is quantum mechanically acceptable, find the acceptable eigen values of the SDE. The eigen energy is given in terms of acceptable eigen value of the SDE.

6. If the new differential equation for $v(\xi)$ is not recognized to be an SDE, then attempt a series solution. This procedure can be followed even if the equation is recognized to be an SDE.
7. Investigate large ξ behavior of $u(\xi)$, if the series for $v(\xi)$ does not terminate. Find the condition on λ, such that the series for $v(\xi)$ terminates and consequently $u(\xi)$ vanishes for $|\xi| \to \infty$. This choice of λ gives the energy eigen value. With this choice of λ, the eigen function is given as $u(\xi) = \phi_\infty(\xi)v(\xi)$ or $u(\xi) = \phi_0(\xi)\phi_\infty(\xi)v(\xi)$, according to the case.
8. Finally normalize the wave function to get the normalized eigen function.

In Sect. 9.4 we will follow these steps to obtain bound states of the one-dimensional harmonic oscillator. In Chap. 11, we will discuss the hydrogen atom, by separating variables and solving the resulting one-dimensional radial Schrödinger equation, according to the above procedure. This procedure will also be followed in Chap. 12 for a number of three-dimensional potentials.

9.4 A Particle in a Harmonic Oscillator Well

As an example of this procedure, we consider the one-dimensional harmonic oscillator problem by solving the coordinate space Schrödinger equation. In Chap. 6, we solved this case by the operator method in the abstract Hilbert space. Thus we solve the quantum mechanical motion of a particle of mass m in a one-dimensional harmonic oscillator potential well

$$V(x) = \frac{1}{2}Kx^2,$$

where K is the stiffness constant for the classical motion, whose restoring force is $-Kx$ directed toward the center. As a function of x the potential is a parabola extending to infinity (Fig. 9.5). Thus neither the classical particle, nor the quantum particle can escape from the infinite well. While the classical particle with a total energy E cannot go beyond the classical turning points $x = \pm a$ (where $E = V(a) = \frac{1}{2}Ka^2$), the corresponding quantum particle with the same total energy can be found even for $|x| > a$.

Importance of the harmonic oscillator in physics

This potential is one of only a few, for which the Schrödinger equation admits an analytic solution. In addition to this, it has a number of general interests from the point of view of physics. The stable state of any physical system corresponds to the minimum of the potential function (in general the potential surface in more than one-dimensional space), both classically and quantum mechanically. An expansion of the potential function about this minimum is a quadratic expression of the displacements, since condition of minimum requires the first derivatives (gradient) to vanish and the second derivative positive (for multi-dimensional case, the matrix of

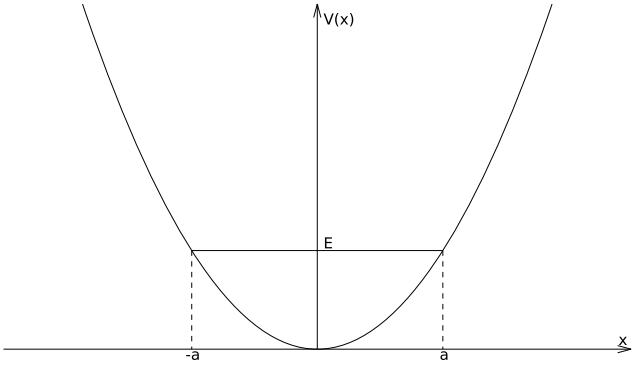


Fig. 9.5 Plot of the harmonic oscillator well. The potential has a minimum value zero at $x = 0$. The classical turning points for an energy E are $\pm a$, such that $E = \frac{1}{2}Ka^2$

the second derivatives should be positive definite). Thus, disregarding the minimum value of the potential $V_{\min}$ (which is permissible in non-relativistic motion, both classical and quantum mechanical), the potential in each of the displacement variables is a harmonic oscillator for small displacements. For motions involving larger displacements in an arbitrary potential, solutions of the harmonic oscillator will be a good starting point for perturbative or variational calculations, since its analytic results are known. In addition these are useful in field quantization (e.g. radiation field quantization).

Solution of the Schrödinger equation

The Schrödinger equation for a particle of mass m and energy E moving in the one-dimensional harmonic oscillator potential is (following Schiff 1968)

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2u(x)}{\mathrm{d}x^2} + \frac{1}{2}Kx^2u(x) = Eu(x). \tag{9.18}$$

The frequency of the classical oscillator is $\omega_c = \sqrt{\frac{K}{m}}$. First, we introduce a dimensionless variable $\xi = \alpha x$ (so that the constant α has dimension $[L^{-1}]$) and multiply through by $-\frac{2m}{\hbar^2\alpha^2}$ to get

$$\frac{\mathrm{d}^2u(\xi)}{\mathrm{d}\xi^2} + \frac{2mE}{\hbar^2\alpha^2}u(\xi) - \frac{mK}{\hbar^2\alpha^4}\xi^2u(\xi) = 0$$

Choose α such that the constant in front of ξ^2 is 1 (so that the equation becomes dimensionless) and call the constant involving E as λ (also a dimensionless quantity)

$$\alpha^4 = \frac{mK}{\hbar^2} \quad \left(\text{hence } \alpha = \sqrt{\frac{m\omega_c}{\hbar}}\right) \qquad \text{and} \qquad \lambda = \frac{2mE}{\hbar^2\alpha^2} = \frac{2E}{\hbar\omega_c}.$$

The differential equation becomes

$$\frac{\mathrm{d}^2 u(\xi)}{\mathrm{d}\xi^2} + (\lambda - \xi^2)u(\xi) = 0$$

Next we investigate the asymptotic behavior. For large $|\xi|$, we can neglect the λ term and inspection shows that $u(\xi) \propto \mathrm{e}^{-\frac{1}{2}\xi^2}$ is a solution to leading power of ξ. Thus we substitute

$$u(\xi) = v(\xi)\mathrm{e}^{-\frac{1}{2}\xi^2}$$

and get (canceling out $\mathrm{e}^{-\frac{1}{2}\xi^2}$, which does not vanish)

$$\frac{\mathrm{d}^2 v(\xi)}{\mathrm{d}\xi^2} - 2\xi\frac{\mathrm{d}v(\xi)}{\mathrm{d}\xi} + (\lambda - 1)v(\xi) = 0$$

or in short hand notation

$$v''(\xi) - 2\xi v'(\xi) + (\lambda - 1)v(\xi) = 0. \tag{9.19}$$

This equation can be recognized as the standard Hermite differential equation (7.37), for the Hermite polynomial $H_n(\xi)$

$$H_n''(\xi) - 2\xi H_n'(\xi) + 2nH_n(\xi) = 0. \tag{9.20}$$

Hence we have $u(\xi) \equiv u_n(\xi) = H_n(\xi)\mathrm{e}^{-\frac{1}{2}\xi^2}$. Here n is a non-negative integer (including zero). This is a requirement for the Hermite function to be a polynomial, so that $u(\xi)$ is finite for $|\xi| \to \infty$. Thus $\lambda = (2n + 1)$, and hence energy eigen values $E_n = (n + \frac{1}{2})\hbar\omega_c$ with $n = 0, 1, 2, 3, \cdots$ become discrete.

However we will solve eq. (9.19) by Frobenius method and see that $u(\xi)$ will diverge as $\xi \to \infty$, unless the series for $v(\xi)$ terminates becoming a polynomial. As in eq. (7.10), we expand $v(\xi)$ about the origin (which is the point of symmetry)

$$v(\xi) = \sum_{i=0}^{\infty} a_i \xi^{s+i} \qquad (a_0 \neq 0), \tag{9.21}$$

where s is the index. Substitution in Eq. (9.19) gives

$$\sum_{i=0}^{\infty} a_i(s+i)(s+i-1)\xi^{s+i-2} - 2\sum_{i=0}^{\infty} a_i(s+i)\xi^{s+i} + (\lambda - 1)\sum_{i=0}^{\infty} a_i\xi^{s+i} = 0$$

Indicial equation is obtained by equating the coefficient of the lowest power of ξ to zero, i.e. coefficient of $\xi^{(s-2)}$

$$a_0 s(s-1) = 0. \qquad \text{Hence } s = 0 \text{ or } 1$$

Equating coefficient of $\xi^{(s-1)}$ to zero, we have

$$a_1 s(s+1) = 0. \qquad \text{Hence for } s = 1 \text{ we must have } a_1 = 0.$$

A general recurrence relation (RR) is obtained by equating the coefficient of a general power of ξ, *viz.* equating the coefficient of $\xi^{(s+i)}$ to zero

$$a_{i+2} = \frac{2(s+i)+1-\lambda}{(s+i+1)(s+i+2)} a_i, \qquad i = 0, 1, 2, \cdots \tag{9.22}$$

For $s = 1, a_1 = 0$, hence all odd subscripted coefficients vanish. For $s = 0$, we get *two* series solutions (one containing only even subscripted coefficients and the other containing only odd subscripted coefficients, as a_{i+2} is related to a_i). However, it can be verified easily that in this case ($s = 0$) the odd subscripted coefficients do not lead to any new linearly independent solution (it gives rise to a term proportional to a_1 times the series for $s = 1$). Thus we can set $a_1 = 0$ (hence $a_1 = a_3 = a_5 = \cdots = 0$) for both $s = 0$ and $s = 1$ and get two linearly independent solutions (as expected) with only even subscripted coefficients. The function $v(\xi)$ is even or odd for $s = 0$ or $s = 1$ respectively. This is consistent with the fact that eq. (9.19) is invariant under parity (see Chap. 8). We can obtain both the linearly independent series solutions using the RR Eq. (9.22) with $s = 0$ and $s = 1$.

We will first investigate what happens, if the series is allowed to become an infinite series. Using Eq. (9.22) the ratio of consecutive terms [$(i+2)$-th and i-th terms] of the series is $\frac{[2(s+i)+1-\lambda]\xi^2}{(s+i+1)(s+i+2)}$. This is true for both $s = 0$ and $s = 1$ and i takes on even integer values only. Thus these two *consecutive terms* (t_m) are really the $(m+1)$-th and m-th term respectively with $m = \frac{i}{2}$ = a non-negative integer. Hence

$$\lim_{m\to\text{large}} \frac{t_{m+1}}{t_m} = \lim_{i\to\text{large}} \frac{t_{\frac{i}{2}+1}}{t_{\frac{i}{2}}} = \lim_{i\to\text{large}} \frac{[2(s+i)+1-\lambda]\xi^2}{(s+i+1)(s+i+2)}.$$
$$\to \frac{2\xi^2}{i} = \frac{\xi^2}{m}.$$

The ratio of the $(m+1)$-th and m-th terms of the series $e^{\xi^2} = \sum_{m=0}^{\infty} \frac{(\xi^2)^m}{m!}$ for large m is

$$\lim_{m\to\text{large}} \frac{m!}{(m+1)!}\xi^2 \to \frac{\xi^2}{m}.$$

Thus, if the series for $v(\xi)$ does not terminate, the tail part of the series will behave as e^{ξ^2} and $u(\xi) = v(\xi)e^{-\frac{1}{2}\xi^2}$ will behave as $e^{\frac{1}{2}\xi^2}$. This diverges for $|\xi| \to \infty$, which is not allowed. On the other hand, if the series for $v(\xi)$ terminates into a polynomial of degree n, the $|\xi| \to \infty$ behavior of $u(\xi)$ will be dominated by $e^{-\frac{1}{2}\xi^2}$ and will vanish. If the series (9.21) terminates at the $(i+s)$-th term, we must have $a_{i+2} = 0$, but $a_i \neq 0$, which requires from Eq. (9.22)

$$\lambda = 2(s+i)+1 = 2n+1, \tag{9.23}$$

where $(s+i) = n$, a non-negative integer. It is the degree of the polynomial. Since i is an even integer, n is even for $s = 0$ and odd for $s = 1$. The quantization condition (9.23) gives the energy eigen value

$$\lambda = \frac{2E}{\hbar\omega_c} = 2n+1. \qquad \text{Hence } E \equiv E_n = (n+\frac{1}{2})\hbar\omega_c, \qquad n = 0, 1, 2, \cdots \tag{9.24}$$

Eigen functions

The eigen function $u_n(\xi) = v_n(\xi)\mathrm{e}^{-\frac{1}{2}\xi^2}$ with $n = 0, 1, 2, \cdots$ can be obtained in terms of a single constant a_0 using RR, Eq. (9.22). The normalization constant (a_0) can be found by the normalization condition. This procedure is useful for small values of n. But this is cumbersome for large n, and is not possible for a general n. For the latter case, we recognize that the differential equation for $v_n(\xi)$ is the Hermite differential equation (9.20) with $v_n(\xi) = H_n(\xi)$ and use standard properties of the Hermite polynomial $H_n(\xi)$ and standard integrals involving them (see Arfken 1966). In the following we outline the process of obtaining such properties and integrals, using the generating function of the Hermite polynomials.

A *generating function* of a standard polynomial function is a function $G(\xi, w)$ of two variables, say ξ and w, whose expansion as a power series in w is specified with the coefficient of w^n involving the standard function of order n. The generating function of the Hermite polynomials is

$$G(\xi, w) = \mathrm{e}^{\xi^2-(w-\xi)^2} = \mathrm{e}^{-w^2+2\xi w} \equiv \sum_{n=0}^{\infty} \frac{H_n(\xi)}{n!} w^n. \tag{9.25}$$

In order to prove that $H_n(\xi)$ appearing in Eq. (9.25) is indeed the Hermite polynomial, we eliminate w from this equation to get a second order differential equation involving ξ only and identify it to be the Hermite differential equation. Differentiate both sides of Eq. (9.25) with respect to (w.r.t) ξ and w and using Eq. (9.25) again

$$\frac{\partial G}{\partial \xi} = 2w\mathrm{e}^{-w^2+2\xi w} = \sum_{n=0}^{\infty} \frac{2H_n(\xi)}{n!} w^{n+1} = \sum_{n=0}^{\infty} \frac{H_n'(\xi)}{n!} w^n$$

$$\frac{\partial G}{\partial w} = (-2w+2\xi)\mathrm{e}^{-w^2+2\xi w} = \sum_{n=0}^{\infty} \frac{H_n(\xi)}{n!}(-2w+2\xi)w^n$$

$$= \sum_{n=1}^{\infty} \frac{H_n(\xi)}{(n-1)!} w^{(n-1)},$$

equating equal powers of w from the two infinite sums in each of the above equations, we have

$$H_n'(\xi) = 2nH_{n-1}(\xi)$$
$$H_{n+1}(\xi) = 2\xi H_n(\xi) - 2nH_{n-1}(\xi). \tag{9.26}$$

These are the recurrence relations of Hermite polynomials. Substituting the first in the second equation, we have

$$H_{n+1}(\xi) = 2\xi H_n(\xi) - H_n'(\xi).$$

Differentiating w.r.t. ξ again and using the first of Eqs. (9.26), we get

$$H_{n+1}'(\xi) = 2(n+1)H_n(\xi) = 2H_n(\xi) + 2\xi H_n'(\xi) - H_n''(\xi)$$
$$\text{or} \quad H_n''(\xi) - 2\xi H_n'(\xi) + 2nH_n(\xi) = 0.$$

This is the standard Hermite differential equation (9.20), proving that the $H_n(\xi)$ appearing in the generating function Eq. (9.25) is indeed the Hermite polynomial of order n.

The n-th eigen function $u_n(\xi)$ of the harmonic oscillator corresponding to the eigen value $E_n = (n + \frac{1}{2})\hbar\omega_c$ is

$$u_n(x) = N_n H_n(\alpha x)\mathrm{e}^{-\frac{1}{2}\alpha^2 x^2}, \tag{9.27}$$

where N_n is the normalization constant, determined through

$$\int_{-\infty}^{\infty} \left|u_n(x)\right|^2 \mathrm{d}x = 1 \;\Rightarrow\; N_n = \left[\frac{1}{\alpha}\int_{-\infty}^{\infty} |H_n(\xi)|^2 \mathrm{e}^{-\xi^2}\mathrm{d}\xi\right]^{-\frac{1}{2}}. \tag{9.28}$$

To obtain explicit expressions for $H_n(\xi)$ and N_n we use the generating function as follows. Differentiate both sides of Eq. (9.25) m times w.r.t. w. All terms on right side with $n < m$ will vanish and all terms with $n \geq m$ will have powers of w, starting with 0. Now setting $w = 0$, only the term $n = m$ will survive. Hence

$$H_m(\xi) = \frac{\partial^m}{\partial w^m}\mathrm{e}^{\xi^2-(w-\xi)^2}\Big|_{w=0} = \mathrm{e}^{\xi^2}\frac{\partial^m}{\partial w^m}\mathrm{e}^{-(w-\xi)^2}\Big|_{w=0}$$
$$= \mathrm{e}^{\xi^2}(-1)^m\frac{\partial^m}{\partial \xi^m}\mathrm{e}^{-(w-\xi)^2}\Big|_{w=0} = \mathrm{e}^{\xi^2}(-1)^m\frac{\partial^m}{\partial \xi^m}\mathrm{e}^{-\xi^2}.$$

The first equality in the second line stems from the fact that for a function of $(w - \xi)$ only, $\frac{\partial}{\partial w} = -\frac{\partial}{\partial \xi}$. Replacing m by n, we get an explicit expression for $H_n(\xi)$

$$H_n(\xi) = (-1)^n\mathrm{e}^{\xi^2}\frac{\mathrm{d}^n}{\mathrm{d}\xi^n}\mathrm{e}^{-\xi^2}. \tag{9.29}$$

Putting $n = 0, 1, 2$ we get explicit expressions for the first three Hermite polynomials

$$H_0(\xi) = 1 \qquad H_1(\xi) = 2\xi \qquad H_2(\xi) = 4\xi^2 - 2.$$

The series solution using Eq. (9.22) yields expressions proportional to these.

Next we will see how integrals involving one or more Hermite polynomials can be evaluated using the generating function (9.25). In particular we first evaluate the normalization integral in Eq. (9.28). Multiply two generating functions $G(\xi, w_1)$, $G(\xi, w_2)$ and $e^{-\xi^2}$ and integrate over ξ from $-\infty$ to ∞

$$\int_{-\infty}^{\infty} e^{-{w_1}^2-{w_2}^2+2\xi(w_1+w_2)-\xi^2} d\xi = \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} \frac{{w_1}^{n_1}{w_2}^{n_2}}{n_1! n_2!} \int_{-\infty}^{\infty} H_{n_1}(\xi) H_{n_2}(\xi) e^{-\xi^2} d\xi. \tag{9.30}$$

The integral on the left side is

$$e^{2w_1w_2} \int_{-\infty}^{\infty} e^{-(\xi-w_1-w_2)^2} d\xi = e^{2w_1w_2}\sqrt{\pi} = \sqrt{\pi}\sum_{m=0}^{\infty} \frac{2^m {w_1}^m {w_2}^m}{m!}. \tag{9.31}$$

Equate the coefficients of ${w_1}^{n_1}{w_2}^{n_2}$ on right sides of the last two equations to obtain the integral over ξ. Note that the powers of w_1 and w_2 in Eq. (9.31) are equal. Hence for $n_1 \neq n_2$, the integral on right side of Eq. (9.30) vanishes. Thus for $n_1 = n_2 = m$

$$\int_{-\infty}^{\infty} H_{n_1}(\xi) H_{n_2}(\xi) e^{-\xi^2} d\xi = \delta_{n_1 n_2}\sqrt{\pi}\, 2^{n_1} n_1!. \tag{9.32}$$

Hence from Eq. (9.28) we have

$$N_n = \left[\frac{\alpha}{\sqrt{\pi}\, 2^n\, n!}\right]^{\frac{1}{2}}. \tag{9.33}$$

Matrix element of a power of x

The matrix element of x^q, $(q = 1, 2, \cdots)$ is [using Eq. (9.27)]

$$\langle n_1 | x^q | n_2 \rangle = \frac{N_{n_1} N_{n_2}}{\alpha^{(q+1)}} \int_{-\infty}^{\infty} H_{n_1}(\xi) H_{n_2}(\xi) e^{-\xi^2} \xi^q d\xi.$$

As before, multiplying $G(\xi, w_1)$, $G(\xi, w_2)$ and $\xi^q e^{-\xi^2}$ (with $q = 1, 2, \cdots$), integrating over ξ from $-\infty$ to ∞ and finally equating coefficients of ${w_1}^{n_1}{w_2}^{n_2}$ we can evaluate matrix elements of x, x^2, $\cdots$. With appropriate change of variables and

renaming, we find that the matrix element of x is the same as Eq. (6.38) obtained by the operator method in Chap. 6. Matrix elements of higher powers of x also agree with those by the matrix multiplication method using matrix elements of x, as discussed in Chap. 6.

Matrix element of powers of momentum

The matrix element of q-th power of momentum ($\hat{p} = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}$) is

$$\langle n_1|\hat{p}^q|n_2\rangle = N_{n_1}N_{n_2}(-i\hbar)^q\alpha^{(q-1)}\int_{-\infty}^{\infty} H_{n_1}(\xi)\mathrm{e}^{-\frac{1}{2}\xi^2}\,\frac{\mathrm{d}^q[H_{n_2}(\xi)\mathrm{e}^{-\frac{1}{2}\xi^2}]}{\mathrm{d}\xi^q}\,\mathrm{d}\xi.$$

q times differentiation of $[H_{n_2}(\xi)\mathrm{e}^{-\frac{1}{2}\xi^2}]$ w.r.t. ξ will yield terms with products of derivatives of $H_{n_2}(\xi)$, powers of ξ and $\mathrm{e}^{-\frac{1}{2}\xi^2}$. Integrals involving such terms can be evaluated by the above method by taking $G(\xi, w_1)\frac{\mathrm{d}^{q_1}G(\xi,w_2)}{\mathrm{d}\xi^{q_1}}\xi^{q_2}\mathrm{e}^{-\xi^2}$. Alternatively, we can use first of the RR of Hermite polynomial, Eq. (9.26) and matrix elements of powers of x (with appropriate coefficients). The matrix element of $\hat{p}$ is

$$\langle n_1|\hat{p}|n_2\rangle = -i\hbar\alpha\Big[\sqrt{\frac{n_1+1}{2}}\delta_{n_2,n_1+1} - \sqrt{\frac{n_1}{2}}\delta_{n_2,n_1-1}\Big]. \tag{9.34}$$

With appropriate change of variables and renaming, this is seen to be the same as Eq. (6.39) of Chap. 6.

The process becomes progressively more cumbrous as q increases. However, we can use completeness of eigen functions

$$\sum_m |m\rangle\langle m| = \hat{1}. \qquad \Rightarrow \qquad \sum_m u_m(x)u_m^*(x') = \delta(x-x') \tag{9.35}$$

and matrix elements of lower power to evaluate those of higher powers, as mentioned after Eq. (6.39). This process is more convenient and is equivalent to matrix multiplication of matrices of lower powers of position or momentum.

All the properties derived in Chap. 6 can be obtained by the coordinate representation method.

Properties of the eigen functions

Here we enumerate certain properties, which we saw as general properties

1. Since the potential $V(x)$ is symmetric about the origin, the eigen functions are either even or odd functions of x. $u_n(x)$ is even for $n = 0, 2, 4, \cdots$ and odd for $n = 1, 3, 5, \cdots$ respectively.
2. The $(n+1)$-th eigen function $u_n(x)$ has n nodes and parity $(-1)^n$ with $n = 0, 1, 2, \cdots$.

Connection between physics and mathematics

1. The domain of regularity (without any intervening singularity) of the Schrödinger equation [as also that of the reduced Hermite differential equation (9.20)] is $-\infty < x < \infty$, with irregular singularities at both the boundaries. The *physical domain* agrees exactly with this domain of mathematical regularity, as discussed in Chap. 8.
2. The asymptotic behavior of the solution $u(\alpha x)$ of the Schrödinger equation is $\mathrm{e}^{-\frac{1}{2}\alpha^2 x^2}$, so that $u(\alpha x) = N_n H_n(\alpha x)\mathrm{e}^{-\frac{1}{2}\alpha^2 x^2}$. All measurable physical quantities are integrals involving bi-linear expressions of $u^*(\alpha x)$ and $u(\alpha x)$. Thus the integrand involves $H_n(\alpha x)$, $H_m(\alpha x)$ and $\mathrm{e}^{-\alpha^2 x^2}$. This agrees with the weight function $\mathrm{e}^{-\alpha^2 x^2}$ of the Hermite differential equation (see Chap 8 and Appendix A).
3. The interval and the weight function make the Hermite differential operator (hence the Hamiltonian) Hermitian, which is a basic requirement according to the fundamental postulates, Chap. 3, Sect. 3.2.
4. Physics requires the wave function to be finite everywhere within the domain. The fact that $x = \pm\infty$ are irregular singular points demands that the series solution of the Hermite differential equation be terminated, which in turn gives rise to energy eigen values. Thus physical requirement and mathematical properties are closely related.

These show the intimate connection between physics of the problem and the inherent mathematics (see Sect. 7.1.4, Chap. 7).

9.5 Wave Packet in a Harmonic Oscillator Well

In Chap. 8, Sect. 8.5 we saw how to obtain a general wave packet satisfying the time-dependent Schrödinger equation in a potential well. Accordingly we can construct a wave packet which moves in a one-dimensional harmonic oscillator well. In this case we have discrete eigen states in the well and according to Eq. (8.20) we construct the wave packet at time t as

$$\psi(x,t) = \sum_{n=0}^{\infty} A_n u_n(x)\mathrm{e}^{-i\frac{E_n}{\hbar}t} = \mathrm{e}^{-\frac{i}{2}\omega_c t}\sum_{n=0}^{\infty} A_n u_n(x)\mathrm{e}^{-in\omega_c t},$$

since $E_n = (n+\frac{1}{2})\hbar\omega_c$. As in Schiff (1968), we assume that initially the wave packet at time $t = 0$ is the ground state ($n = 0$) of the time-independent Schrödinger equation and is localized at $x = a$. The expansion coefficient A_n can be found from this initial condition [from Eqs. (9.27) and (9.33) with $n = 0$ and $H_0 = 1$]

$$\psi(x,0) = u_0(x-a) = \sqrt{\frac{\alpha}{\sqrt{\pi}}}\, e^{-\frac{1}{2}\alpha^2(x-a)^2},$$

and using orthonormality of the eigen functions $\{u_n(x),\ n = 0, 1, 2, \cdots\}$. Substitution of A_n back in the general form of $\psi(x, t)$ gives $\psi(x, t)$ after some lengthy calculations (see Schiff 1968). From this we have the probability density

$$|\psi(x, t)|^2 = \frac{\alpha}{\sqrt{\pi}} e^{-\alpha^2(x - a\cos\omega_c t)^2}.$$

Comparing this with the probability density of the initial state

$$|u_0(x - a)|^2 = \frac{\alpha}{\sqrt{\pi}} e^{-\alpha^2(x-a)^2},$$

we see that $\psi(x, t)$ represents a wave packet, which at time t has a Gaussian envelope (profile) centered at $a\cos\omega_c t$. Hence as time increases from $t = 0$, the center of the packet moves like the position of the corresponding classical particle, $x_c = a\cos\omega_c t$. Thus the average motion of the packet follows the classical motion in accordance with Ehrenfest's theorem (Chap. 8, Sect. 8.6). Note that the wave packet does not spread with time as it follows the classical motion. This is unlike the spreading for a general wave packet discussed in Chap. 8, Sect. 8.5 and is due to the initial condition *viz.* at $t = 0$, the wave packet coincides with the ground state in the well, which is a minimum uncertainty state (see Chap. 6, Sect. 6.5).

9.6 Potential with a Dirac Delta Function

A Dirac delta function at a point makes the potential non-analytic at that point. Although the wave function is continuous, its derivative is discontinuous at that point. It requires a modified treatment, as we see for a simple example.

Consider the motion of a particle in the potential

$$\begin{aligned} V(x) &= +\infty \quad \text{for } x < -a \text{ and } x > a \\ &= C\delta(x) \quad \text{for } -a \le x \le a. \end{aligned}$$

This is an infinite square well of width $2a$ (Fig. 9.1), with a delta function of strength C at its center. The wave function vanishes in $|x| > a$ The Schrödinger equation in $-a \le x \le a$ is

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2u}{\mathrm{d}x^2} + C\delta(x)u(x) = Eu(x). \tag{9.36}$$

For $x \neq 0$, we have

$$\frac{\mathrm{d}^2u}{\mathrm{d}x^2} + \alpha^2 u(x) = 0, \qquad \left(\alpha = \sqrt{\frac{2mE}{\hbar^2}}\right). \tag{9.37}$$

Let us name the regions $-a \le x \le 0$ and $0 \le x \le a$ as Regions I and II respectively. Solution of Eq. (9.37) is a linear combination of $\sin \alpha x$ and $\cos \alpha x$. Now $u(x)$ must vanish at $x = \pm a$. We can easily see that the correct linear combination can be written in the two regions as

$$\begin{aligned} u_I(x) &= A \sin \alpha(x + a) \\ u_{II}(x) &= B \sin \alpha(x - a). \end{aligned}$$

Continuity of wave function at $x = 0$ requires $B = -A$. Hence

$$\begin{aligned} u_I(x) &= A \sin \alpha(x + a) \\ u_{II}(x) &= -A \sin \alpha(x - a). \end{aligned} \tag{9.38}$$

To get the discontinuity across the delta function, integrate Eq. (9.36) from $x = 0 - \epsilon$ to $x = 0 + \epsilon$, where ϵ is small

$$\int_{x=-\epsilon}^{x=+\epsilon} \frac{\mathrm{d}^2 u}{\mathrm{d}x^2} \mathrm{d}x - \frac{2mC}{\hbar^2} \int_{-\epsilon}^{\epsilon} \delta(x) u(x) \mathrm{d}x = -\frac{2mE}{\hbar^2} \int_{-\epsilon}^{\epsilon} u(x) \mathrm{d}x,$$

and then take the limit $\epsilon \to 0$. Right side vanishes due to continuity of $u(x)$ at $x = 0$. Hence we get

$$\lim_{\epsilon \to 0} \left[\left. \frac{\mathrm{d}u}{\mathrm{d}x} \right|_{x=\epsilon} - \left. \frac{\mathrm{d}u}{\mathrm{d}x} \right|_{x=-\epsilon} \right] = \frac{2mC}{\hbar^2} u(0). \tag{9.39}$$

Substituting for u_I and u_{II} from Eq. (9.38), we get

$$\tan \alpha a = -\frac{\hbar^2}{mCa} \alpha a.$$

Let $\xi = \alpha a$. The transcendental equation for energy becomes

$$\tan \xi = -C_1 \xi, \quad \text{where } C_1 = \frac{\hbar^2}{mCa}. \tag{9.40}$$

This equation can be solved graphically or numerically as in the case of a finite square well. For a graphical solution, we put $\eta = \tan \xi$. Then from Eq. (9.40), we have $\eta = -C_1 \xi$. In Fig. 9.6 we plot these two curves for two different values of $C_1 = 0.1$ and $C_1 = 1$. The intersections of $\eta = \tan \xi$ and the straight line gives a solution of Eq. (9.40). Energy eigen value corresponding to a solution ξ_n is given by Eq. (9.37)

$$E_n = \frac{\hbar^2}{2ma^2} \xi_n^2.$$

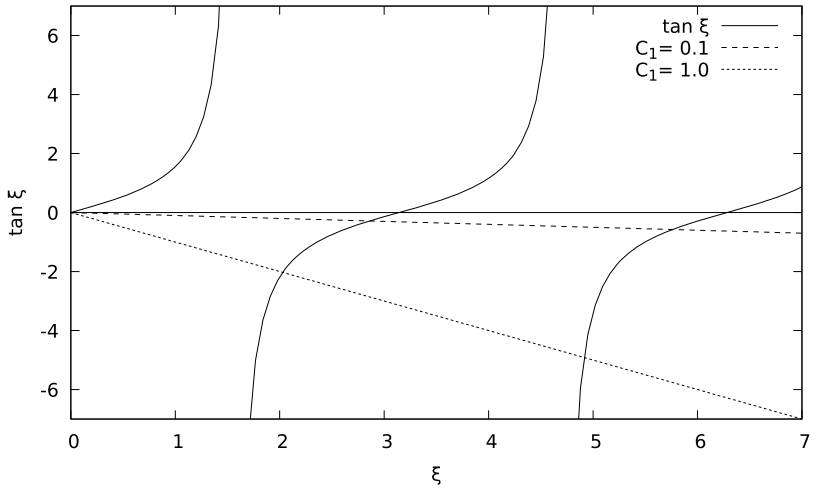


Fig. 9.6 Graphical solution of Eq. (9.40) for $C_1 = 0.1$ and $C_1 = 1$

From Fig. 9.6, we can see that for small C_1 the solution ξ_n is close to, but slightly less than $n\pi$, with $n = 1, 2, \cdots$. On the other hand for large C_1, the solutions are close to, but slightly larger than $(n - \frac{1}{2})\pi$, with $n = 1, 2, \cdots$. In the limit $C = 0$, i.e. $C_1 \to \infty$, we have $\xi_n = n\frac{\pi}{2}$ with $n = 1, 3, \cdots$. This gives $E_n = \frac{\hbar^2\pi^2 n^2}{8ma^2}$, in agreement with Eq. (9.5) for even parity states of the infinite square well. Note that the wave function (9.38) has even parity. For the odd parity states, $u(0)$ must vanish. Then Eq. (9.39) shows that the slopes from the left and right sides of the origin match. Thus $u(x)$ is analytic about $x = 0$. From Eq. (9.38), the condition $u(0) = 0$ gives $\alpha a = n\pi$ with $n = 1, 2, \cdots$, i.e. $\alpha a = n\frac{\pi}{2}$ with $n = 2, 4, \cdots$, which gives $E_n = \frac{\hbar^2\pi^2 n^2}{8ma^2}$ with $n = 2, 4, \cdots$, in complete agreement with the odd parity states of the infinite square well.

Note that the requirement of $u(0) = 0$ for odd parity states makes the quantum particle not to sense the delta function at $x = 0$.

9.7 Quasi-bound State in a δ-Function Barrier

Consider a particle of mass m trapped between an infinite wall and a δ-function in a potential (Fig. 9.7)

$$\begin{aligned} V(x) &= +\infty \qquad \text{for } x < 0 \\ &= C\delta(x - a) \quad \text{for } x \geq 0. \qquad (C \text{ is a constant } > 0) \end{aligned}$$

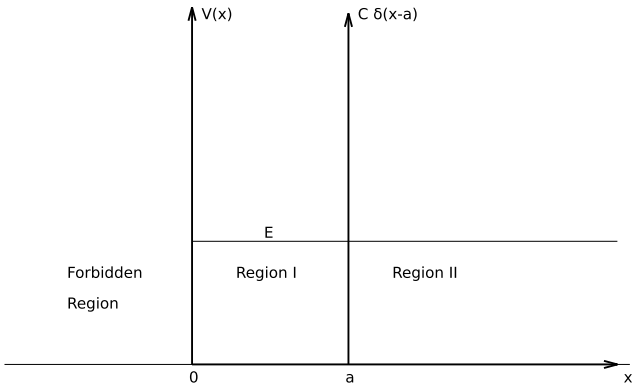


Fig. 9.7 One-dimensional δ-function potential $C\delta(x-a)$ at $x=a$. The region $x<0$ is forbidden

The wave function in the forbidden region $x<0$ is $u(x)=0$ and the Schrödinger equation for $x\geq 0$ is

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2u(x)}{\mathrm{d}x^2}+C\delta(x-a)u(x)=Eu(x). \tag{9.41}$$

For $x>0$ and $x\neq a$

$$\frac{\mathrm{d}^2u}{\mathrm{d}x^2}+\alpha^2u=0,\quad \left(\alpha=\sqrt{\frac{2mE}{\hbar^2}}\right).$$

In Region I (see Fig. 9.7), $u_I(0)=0$, hence

$$u_I(x)=A\sin\alpha x,$$

and in Region II, since u_{II} represents a transmitted wave traveling to the right

$$u_{II}(x)=F\mathrm{e}^{i\alpha x}.$$

Continuity of wave function at $x=a$ is

$$A\sin\alpha a=F\mathrm{e}^{i\alpha a}. \tag{9.42}$$

Discontinuity of the slopes at $x=a$ can be found by integrating Eq. (9.41) from $x=a-\epsilon$ to $x=a+\epsilon$ and then taking the limit $\epsilon\to 0$ [note that the term coming from right side of Eq. (9.41) vanishes due to continuity of u]

$$\left.\frac{\mathrm{d}u_{II}}{\mathrm{d}x}\right|_{x=a}-\left.\frac{du_I}{\mathrm{d}x}\right|_{x=a}=\frac{2mC}{\hbar^2}u_I(a).$$

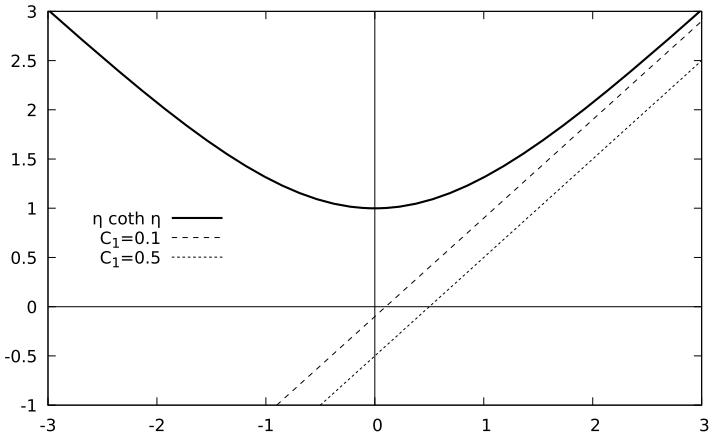


Fig. 9.8 Plot of $\eta \coth \eta$ and $-C_1 + \eta$ (straight lines) as functions of η, for $C_1 = 0.1$ and $C_1 = 0.5$

Using expressions for u_I and u_{II}, we have

$$i\alpha F e^{i\alpha a} - A\alpha \cos \alpha a = \frac{2mC}{\hbar^2} A \sin \alpha a.$$

Using Eq.(9.42), the first term on left is $i\alpha A \sin \alpha a$. Hence an equation for energy becomes

$$i\alpha \sin \alpha a - \alpha \cos \alpha a = \frac{2mC}{\hbar^2} \sin \alpha a. \tag{9.43}$$

Introducing a dimensionless variable ξ and a dimensionless real positive constant C_1 through

$$\xi = \alpha a \quad \text{and} \quad C_1 = \frac{2mCa}{\hbar^2},$$

in Eq. (9.43), we have a dimensionless energy equation

$$\xi \cot \xi = -C_1 + i\xi$$

Further substituting $\xi = -i\eta$, we have $\cot \xi = i \coth \eta$. Hence

$$\eta \coth \eta = -C_1 + \eta. \tag{9.44}$$

We now have a transcendental equation with real coefficients. To solve it graphically we plot the curve $\eta \coth \eta$ and the straight line $-C_1 + \eta$ in Fig. 9.8 for $C_1 = 0.1$ and $C_1 = 0.5$. We see that neither of the straight lines intersect the curve. In fact the straight line for $C_1 = 0$ is the asymptotic tangent to the curve $\eta \coth \eta$ for large

positive η. Hence for positive C_1 (as required for the repulsive δ-function) there is no real solution. Thus η and hence α and E must be complex.[1] Imaginary part of E must be negative for the probability density to decrease with time, as the particle leaks outwards with increasing time. Constants appearing in $u(x,t)$ are also complex, since they are related by Eq. (9.42). Such a barrier traps the particle in a quasi-bound state. The total probability decreases with time (as E is complex) and corresponds to tunneling through the barrier (see Chap. 8, Sect. 8.2).

9.8 Problems

1. Solve the quantum mechanical problem of a particle of mass m moving in the one-dimensional potential

$$\begin{aligned} V(x) &= +\infty \quad \text{for } x < 0 \\ &= 0 \quad \text{for } 0 \le x \le 2a \\ &= +\infty \quad \text{for } x > 2a \end{aligned}$$

   Show that the energy eigen values are the same as in Sect. 9.1. Show also that the change in the eigen functions is as expected.
2. Solve the problem of a particle in a finite square well (Sect. 9.2) without using the symmetry and show that the final results are the same.
3. Solve the three-dimensional problem of a particle in a cubical box with rigid walls. Show that the Schrödinger equation readily separates in Cartesian coordinate system. Then use the result of Sect. 9.1 of this chapter to obtain the energy eigen values and their eigen functions. Study the degeneracy of a particular eigen value.
4. Study the time evolution of a state which is a linear combination (with equal coefficients) of the first three eigen functions of the one-dimensional harmonic oscillator. Calculate position probability density as a function of x and t.
5. Calculate the energy levels and eigen functions for a particle in a one-dimensional "half-harmonic oscillator" potential well

$$\begin{aligned} V(x) &= \frac{1}{2}m\omega^2 x^2 \quad \text{for } x > 0 \\ &= \infty \quad \text{for } x \le 0. \end{aligned}$$

   Sketch the potential.

   [Hint: The Schrödinger equation for $x > 0$ is the same as that of the "full" harmonic oscillator well.]

[1] We can see that ξ and hence α are complex from the ξ-equation directly. From Eq. (9.44), we demonstrate that no real solution exists.

6. Solve the quantum mechanical problem of a particle of mass m moving in the one-dimensional potential of Fig. 8.10

$$\begin{aligned} V(x) &= +\infty, \quad x < 0 \\ &= 0, \quad 0 \le x \le a \\ &= +V_o, \quad a < x \le b \\ &= 0, \quad b < x < \infty \end{aligned}$$

with V_0, a, b positive and $b > a$.
(a) Will there be bound states with energy $0 < E < V_0$? What is the nature of such states?
(b) Set up an equation for energy E in the range $0 < E < V_0$. Show that solution for E is complex. Obtain the time dependence for the probability density in terms of the imaginary part of E.

References

Arfken, G.: Mathematical Methods for Physicists. Academic Press, New York (1966)
Schiff, L.I.: Quantum Mechanics, 3rd edn. McGraw-Hill Book Company Inc., Singapore (1968)

Chapter 10
Three-Dimensional Problem: Spherically Symmetric Potential and Orbital Angular Momentum

Abstract Connection between a spherically symmetric potential and conserved orbital angular momentum is discussed. Eigen function of orbital angular momentum operator in spherical polar angles is obtained as spherical harmonics. Physical interpretation of the separated radial equation has been provided.

Keywords Orbital angular momentum · Its conservation in spherically symmetric potentials · Eigen values and eigen functions · Separated radial equation

The Schrödinger equation for bound states of a particle of mass μ moving in a three-dimensional potential $V(\vec{r})$ is

$$-\frac{\hbar^2}{2\mu}\nabla_{\vec{r}}^2\, u(\vec{r}) + V(\vec{r})\, u(\vec{r}) = E\, u(\vec{r}). \tag{10.1}$$

An important class of potentials consists of spherically symmetric potentials: $V(\vec{r}) = V(r)$ and is independent of the orientation of $\vec{r}$. Such potentials are also called *central force potential*, since the force is along $\vec{r}$ and is directed toward or away from the origin (center). For such potentials, the Schrödinger equation is separable in spherical polar coordinates $\vec{r} = (r, \theta, \phi)$, in which

$$\nabla_{\vec{r}}^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \tag{10.2}$$

In the following we will see how Eq. (10.1) for spherically symmetric potentials can be separated in polar coordinates, using orbital angular momentum.

10.1 Connection with Orbital Angular Momentum

For a spherically symmetric potential, the Hamiltonian is invariant under 3D rotations. Hence orbital angular momentum is conserved. Consequently, its quantum mechanically allowed magnitude and one component (see below) lead to good quan-

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5_10

tum numbers. Classical expression of orbital angular momentum is $\vec{L} = \vec{r} \times \vec{p}$. Corresponding quantum mechanical operator is (Schiff 1955; Merzbacher 1965)

$$\hat{\vec{L}} = \hat{\vec{r}} \times \hat{\vec{p}} \tag{10.3}$$

$$= (-i\hbar)\, \vec{r} \times \vec{\nabla}_{\vec{r}} \tag{10.4}$$

From the Cartesian components of Eq. (10.3), we get

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \quad \hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z, \quad \hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_z. \tag{10.5}$$

The Cartesian components of $\vec{r}$ are

$$\begin{aligned} x &= r\ \sin\theta\ \cos\phi \\ y &= r\ \sin\theta\ \sin\phi \\ z &= r\ \cos\theta. \end{aligned}$$

The gradient operator in spherical polar coordinates (with unit vectors $\hat{i}_r$. $\hat{i}_\theta$ and $\hat{i}_\phi$) is (Merzbacher 1965)

$$\vec{\nabla}_{\vec{r}} = \hat{i}_r \frac{\partial}{\partial r} + \hat{i}_\theta \frac{1}{r}\frac{\partial}{\partial \theta} + \hat{i}_\phi \frac{1}{r\sin\theta}\frac{\partial}{\partial \phi},$$

and $\hat{\vec{L}}$ is given by

$$\hat{\vec{L}} = -i\hbar\Big[-\hat{i}_\theta \frac{1}{\sin\theta}\frac{\partial}{\partial \phi} + \hat{i}_\phi \frac{\partial}{\partial \theta}\Big].$$

Now using Eq. (10.5) and the fundamental commutation relations between the components of $\hat{\vec{r}}$ and $\hat{\vec{p}}$, we have the well known commutation relations between components of $\hat{\vec{L}}$

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z, \quad [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \quad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y. \tag{10.6}$$

The square of orbital angular momentum is

$$\hat{L}^2 = \hat{\vec{L}}.\hat{\vec{L}} = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$$

Using the above relations we have (Merzbacher 1965)

$$\hat{L}^2 = -\hbar^2\Big[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\Big(\sin\theta\frac{\partial}{\partial\theta}\Big) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\Big]. \tag{10.7}$$

Substituting in Eq. (10.2), we get

$$\nabla_{\vec{r}}^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{\hat{L}^2}{\hbar^2 r^2}. \tag{10.8}$$

Using this, the Schrödinger equation (10.1) for $V(\vec{r}) = V(r)$ becomes

$$-\frac{\hbar^2}{2\mu}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)\right]u(\vec{r}) + \left[V(r) + \frac{\hat{L}^2}{2\mu r^2}\right]u(\vec{r}) = E\ u(\vec{r}). \tag{10.9}$$

In this only $\hat{L}^2$ involves (θ, ϕ) and no r [see Eq. (10.7)]. Hence we can separate variables r and (θ, ϕ) by writing

$$u(\vec{r}) = R(r)Y(\theta, \phi),$$

substitute in Eq. (10.9) and divide by RY to get an equation whose one side is a function of r only and the other side a function of $(\theta.\phi)$ only. Since these relations must be valid for all values of r and (θ, ϕ), each side must be a constant [independent of (r, θ, ϕ)], say $\lambda\hbar^2$ (we include $\hbar^2$ so that λ is dimensionless). Then we have two separate equations one involving (θ, ϕ) only and the other involving r only

$$\hat{L}^2 Y(\theta, \phi) = \lambda\hbar^2 Y(\theta, \phi) \tag{10.10}$$

$$-\frac{\hbar^2}{2\mu}\left[\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right)\right]R(r) + \left[V(r) + \frac{\lambda\hbar^2}{2\mu r^2}\right]R(r) = E\ R(r). \tag{10.11}$$

Thus energy eigen value (which depends on λ) is determined by Eq. (10.11). However the value of λ is to be determined by the solution of Eq. (10.10).

10.2 Eigen Solution of Orbital Angular Momentum

All Cartesian components of $\hat{\vec{L}}$, as also its square are Hermitian operators and are measurable (see Chap. 3). However, Eq. (10.6) shows that no two components commute. Hence no two components are simultaneously measurable (see Chap. 3). But we can easily verify that

$$[\hat{\vec{L}}^2, \hat{L}_x] = [\hat{\vec{L}}^2, \hat{L}_y] = [\hat{\vec{L}}^2, \hat{L}_z] = 0, \tag{10.12}$$

Hence together with $\hat{\vec{L}}^2$, *any one component* can be simultaneously specified. Conventionally, it is the z-component $\hat{L}_z$. The simultaneous eigen vector of $\hat{\vec{L}}^2$ and $\hat{L}_z$ (corresponding to eigen values $\hbar^2 l(l+1)$ and $m\hbar$ respectively) is denoted by $|lm\rangle$ and satisfies

$$\hat{\vec{L}}^2|lm\rangle = l(l+1)\hbar^2|lm\rangle,$$
$$\hat{L}_z|lm\rangle = m\hbar|lm\rangle.$$

As we see shortly, projection of $|lm\rangle$ on to $\vec{r}$-space is called the spherical harmonics

$$Y_{lm}(\theta, \phi) = \langle \vec{r} | lm \rangle$$

(note that $\hat{\vec{L}}^2$ does not involve r), and $\hat{L}_z$ depends only on ϕ, satisfying an eigen value equation

$$\hat{L}_z \Phi(\phi) = m\hbar \Phi(\phi), \tag{10.13}$$

where $m\hbar$ is the eigen value ($\hbar$ is included, so that m is dimensionless) and $\Phi(\phi)$ is the corresponding eigen function. Frequently, m is called the *magnetic quantum number*, because the z-component of the magnetic moment of a particle of charge e and mass μ is $\frac{e}{2\mu}\hat{L}_z$. The eigen value equation for $\hat{\vec{L}}^2$ is Eq. (10.10), to which we will come back later. In terms of spherical polar coordinates $\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi}$ and Eq. (10.13) becomes

$$-i\frac{d\Phi(\phi)}{d\phi} = m\ \Phi(\phi). \tag{10.14}$$

Its solution is

$$\Phi(\phi) = C_\phi\ \mathrm{e}^{im\phi},$$

where C_ϕ is the normalization constant. ϕ is a cyclic variable and its principal interval is $0 \le \phi < 2\pi$. Hence for continuity of the wave function $\Phi(\phi)$, we must have $\Phi(2\pi + \phi) = \Phi(\phi)$, which demands $\mathrm{e}^{im2\pi} = 1$. This is satisfied if m is an integer, which can be positive, negative or zero. Normalizing $\Phi(\phi)$ according to $\int_0^{2\pi} |\Phi(\phi)|^2 d\phi = 1$, we get $C_\phi = \frac{1}{\sqrt{2\pi}}$. Finally we have

$$\Phi(\phi) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{im\phi}, \quad m = \cdots, -2, -1, 0, 1, 2, \cdots \tag{10.15}$$

It can be easily verified that eigen functions belonging to different eigen values m and m' are orthonormal

$$\int_0^{2\pi} \Phi_m^*(\phi)\Phi_{m'}(\phi)d\phi = \delta_{m,m'}.$$

We next return to the eigen value equation for $\hat{\vec{L}}^2$, Eq. (10.10). From Eq. (10.7) it is seen that ϕ is easily separated by writing $Y(\theta, \phi) = \Theta(\theta)\Phi(\phi)$ and multiplying by $\frac{\sin^2\theta}{\Theta\Phi}$. We have

$$\frac{\sin\theta}{\Theta}\frac{d}{d\theta}\left(\sin\theta\frac{\mathrm{d}\Theta}{\mathrm{d}\theta}\right) + \lambda\sin^2\theta = -\frac{1}{\Phi}\frac{\mathrm{d}^2\Phi}{\mathrm{d}\phi^2}.$$

Since left side is a function of θ only and right side is a function of ϕ only, each side must be equal to a constant. Now, as $\hat{L}_z$ can be specified simultaneously with the eigen value λ of $\hat{L}^2$, we can choose $\Phi(\phi)$ to be the eigen function of $\hat{L}_z$ given by Eqs. (10.13) and (10.15) and we have

$$-\frac{1}{\Phi}\frac{d^2\Phi}{d\phi^2} = m^2.$$

Substituting this, the θ equation becomes

$$\left[\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\frac{\mathrm{d}}{\mathrm{d}\theta}\right) + \left(\lambda - \frac{m^2}{\sin^2\theta}\right)\right]\Theta(\theta) = 0.$$

Changing variable θ to w by introducing $w = \cos\theta$ and writing $\Theta(\theta) = P_\lambda^m(w)$ (we specify the eigen function P by the eigen values λ and m), we get

$$\left[\frac{\mathrm{d}}{\mathrm{d}w}\left((1-w^2)\frac{\mathrm{d}}{\mathrm{d}w}\right) + \left(\lambda - \frac{m^2}{1-w^2}\right)\right]P_\lambda^m(w) = 0. \tag{10.16}$$

Comparing with Eq. (7.45), we identify this equation as the associated Legendre differential equation. As a special case, consider $m = 0$ corresponding to axially symmetric solutions (we write $P_\lambda(w)$ for $P_\lambda^0(w)$)

$$\left[\frac{\mathrm{d}}{\mathrm{d}w}\left((1-w^2)\frac{\mathrm{d}}{\mathrm{d}w}\right) + \lambda\right]P_\lambda(w) = 0. \tag{10.17}$$

Comparing with Eq. (7.41), we identify this as the Legendre differential equation. For $m = 0$ we suppress the superscript m and the function P_λ is called the Legendre function. As in Chap. 7, Sect. 7.1.3, we try a series solution

$$P_\lambda(w) = \sum_{i=0} a_i w^{i+s} \quad a_0 \neq 0.$$

Substitution in Eq. (10.17) gives the solution of the indicial equation: $s = 0$ and 1. The recurrence relation (RR) is

$$a_{i+2} = \frac{(i+s)(i+s+1) - \lambda}{(i+s+1)(i+s+2)}a_i. \tag{10.18}$$

Equation (10.17) is invariant under parity $w \to -w$. Hence $P_\lambda(w)$ must be either even or odd, in agreement with the above RR which relates a_{i+2} with a_i. So we can choose $a_1 = 0$, making all the odd subscripted coefficients vanish. Then $s = 0$ and $s = 1$ respectively give the even and odd functions. Now if the series does not terminate, the ratio of successive terms of tail part of the series for $w = \pm 1$ is [note that i takes the values $0, 2, 4, \cdots$, hence the successive terms are the $(\frac{i}{2}+1)$-th and $(\frac{i}{2})$-th terms]

$$\lim_{i\to \text{ large}} \frac{t_{\frac{i}{2}+1}}{t_{\frac{i}{2}}} = \lim_{i\to \text{ large}} \frac{a_{i+2}}{a_i} w^2 = \frac{i}{i+2} = \frac{\frac{i}{2}}{\frac{i}{2}+1}.$$

This ratio is the same as that of the series $\sum_j \frac{1}{j}$, which diverges. Hence if the series for $P_\lambda(w)$ does not terminate, it diverges for $w = \pm 1$, i.e. for $\theta = 0$ or π. This is not permissible, as it is an eigen function of an observable. So the series must terminate to a polynomial. From Eq. (10.18), we see that this is possible if $\lambda = l(l+1)$, where $l = i + s$. Then $a_{i+2} = 0$, while $a_i \neq 0$, and all higher coefficients vanish. Now $l = i + s$ is a positive integer or zero (even for $s = 0$ and odd for $s = 1$, since i is always even). This gives the quantization of orbital angular momentum squared

$$\text{Eigen value of } \hat{\vec{L}}^2 = l(l+1)\hbar^2, \quad l = 0, 1, 2, \cdots, \tag{10.19}$$

where l is called the *orbital angular momentum quantum number*. The corresponding solution of Eq. (10.17) denoted by $P_l(w)$ is a polynomial of degree l, has parity $(-1)^l$ and is called the *Legendre polynomial*. Besides solving the differential equation by the series method, the Legendre polynomial and all its properties can be obtained from its generating function (Chattopadhyay 2006)

$$(1 - 2wv + v^2)^{-\frac{1}{2}} = \sum_{l=0}^{\infty} P_l(w) v^l. \tag{10.20}$$

Another useful form is the *Rodrigues' formula* (Chattopadhyay 2006)

$$P_l(w) = \frac{1}{2^l . l!} \frac{\mathrm{d}^l}{\mathrm{d}w^l} (w^2 - 1)^l. \tag{10.21}$$

Explicit expressions of $P_l(w)$ for small l can be obtained by differentiating Eq. (10.20) l times w.r.t. v and then setting $v = 0$. Equation (10.21) gives $P_l(w)$ directly. The first few are

$$\begin{aligned} P_0(w) &= 1 \\ P_1(w) &= w \\ P_2(w) &= \frac{1}{2}(3w^2 - 1) \\ P_3(w) &= \frac{1}{2}(5w^3 - 3w). \end{aligned}$$

Either the generating function or the Rodrigues' formula can be used to get other important properties, e.g., the orthonormality relation

$$\int_{-1}^{1} P_l(w) P_{l'}(w) \mathrm{d}w = \frac{2}{2l+1} \delta_{l,l'}. \tag{10.22}$$

We have seen that l must be a positive integer or zero. Next we come back to Eq. (10.16) for $m \neq 0$. Note that this equation remains unchanged when m changes sign. We can verify that differentiating Legendre equation (10.17) $|m|$ times w.r.t. w, and defining *associated Legendre functions* as (Schiff 1955)

$$P_l^m(w) = (1-w^2)^{|m|/2} \frac{\mathrm{d}^{|m|} P_l(w)}{\mathrm{d}w^{|m|}}, \tag{10.23}$$

we get associated Legendre differential equation (10.16):

$$\left[\frac{\mathrm{d}}{\mathrm{d}w}\left((1-w^2)\frac{\mathrm{d}}{\mathrm{d}w}\right) + \left(l(l+1) - \frac{m^2}{1-w^2}\right)\right] P_l^m(w) = 0. \tag{10.24}$$

Comparing with Eq. (10.16) we see that $P_l^m(w)$ defined by Eq. (10.23) is the solution of the associated legendre equation, with $\lambda = l(l+1), l = 0, 1, 2, \cdots$. Furthermore, using the Rodrigues' formula for Legendre polynomials Eq. (10.21) in Eq. (10.23), we get Rodrigues' formula for associated Legendre functions as

$$P_l^m(w) = (1-w^2)^{|m|/2} \frac{\mathrm{d}^{|m|} P_l(w)}{\mathrm{d}w^{|m|}} = \frac{1}{2^l.l!}(1-w^2)^{|m|/2} \frac{\mathrm{d}^{l+|m|}}{\mathrm{d}w^{l+|m|}}(w^2-1)^l. \tag{10.25}$$

Note that for $m \neq$ an even integer, these are no more polynomials and in general $P_l^m(w)$ is called associated Legendre *function*. Since $(w^2-1)^l$ is a polynomial in w of degree $2l$, its differentiation more than $2l$ times w.r.t. w gives zero. Hence, in order that $P_l^m(w)$ be a part of the full *non-trivial* wave function, we must have $|m| \leq l$, , i.e. $-l \leq m \leq l$. Thus using Eqs. (10.15) and (10.19), complete selection rule for m becomes: $m = -l, -l+1, \cdots, l-1, l$ where l is a non-negative integer.

Mathematical note: Definition (10.25) has been adopted in Schiff (1955). Sometimes, extending Eq. (10.25), $P_l^m(w)$ is defined as (see Merzbacher 1965; Chattopadhyay 2006; Roy and Nigam 1967).

$$P_l^m(w) = \frac{1}{2^l.l!}(1-w^2)^{m/2} \frac{\mathrm{d}^{l+m}}{\mathrm{d}w^{l+m}}(w^2-1)^l,$$

for both positive and negative m. Since $(w^2-1)^l$ is a polynomial in w of degree $2l$, differentiation of it $(l+m)$-times w.r.t. w permits m to be negative $\geq -l$. On the other hand, for positive $m > l$, right side of the last equation vanishes identically. Hence we can extend the definition of associated Legendre functions, $P_l^m(w)$ to negative values of m such that $|m| \leq l$, as in the last equation. Since Eq. (10.16) remains unchanged when m is replaced by $-m$, $P_l^{|m|}(w)$ can be used as its solution. In fact $P_l^m(w)$ and $P_l^{-m}(w)$ are proportional (see Chattopadhyay 2006).

Equation (10.24) is an eigen value equation with eigen value $l(l+1)$ for a fixed m. Hence its solutions corresponding to different l, but the same m are orthogonal

$$\int_{-1}^{1} P_l^m(w) P_{l'}^m(w) \mathrm{d}w = \frac{2}{2l+1} \frac{(l+|m|)!}{(l-|m|)!} \delta_{l,l'}. \tag{10.26}$$

This relation is valid for both positive and negative m. For $m = 0$, we have $P_l^0(w) = P_l(w)$ and Eq. (10.26) reduces to Eq. (10.22).

Spherical harmonics

The complete solution of Eq. (10.10) is obtained from Eqs. (10.15) and (10.25) with $w = \cos\theta$, as

$$Y_{lm}(\theta, \phi) = N_{lm}\ P_l^m(\cos\theta)\ \mathrm{e}^{im\phi},$$

where N_{lm} is the combined normalization constant for both $P_l^m(w)$ and $\Phi(\phi)$ and is obtained from Eqs. (10.26) and (10.15). Normalized $Y_{lm}(\theta, \phi)$ is called *spherical harmonics* and is given by

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-|m|)!}{(l+|m|)!}}\ P_l^m(\cos\theta)\ \mathrm{e}^{im\phi}, \tag{10.27}$$

where $P_l^m(w)$ is defined by Eq. (10.25). $Y_{lm}(\theta, \phi)$ is the normalized simultaneous eigen function of $\hat{L}^2$ and $\hat{L}_z$ corresponding to eigen values $l(l+1)\hbar^2$ and $m\hbar$ respectively

$$\begin{aligned} \hat{L}^2 Y_{lm}(\theta, \phi) &= l(l+1)\hbar^2 Y_{lm}(\theta, \phi) \\ \hat{L}_z Y_{lm}(\theta, \phi) &= m\hbar Y_{lm}(\theta, \phi), \end{aligned} \tag{10.28}$$

where l is a non-negative integer and $m = -l, -l+1, \cdots, l-1, l$. Under parity operation, $\vec{r} \to -\vec{r}$, i.e. $(r, \theta, \phi) \to (r, \pi - \theta, \pi + \phi)$, spherical harmonics $Y_{lm}(\theta, \phi)$ transforms as

$$Y_{lm}(\theta, \phi) \underset{parity}{\longrightarrow} Y_{lm}(\pi - \theta, \pi + \phi) = (-1)^l Y_{lm}(\theta, \phi). \tag{10.29}$$

Orthonormality of spherical harmonics is given by

$$\int_{\theta=0}^{\pi} \sin\theta \mathrm{d}\theta \int_{\phi=0}^{2\pi} \mathrm{d}\phi\ Y_{lm}^*(\theta.\phi) Y_{l',m'}(\theta, \phi) = \delta_{l,l'} \delta_{m,m'}.$$

Condon and Shortley phase
The phase of spherical harmnics is not unique. The phase in Eq. (10.27) corresponds to that of Schiff (1955). A different phase, known as *Condon and Shortley phase* is often used, which has an additional phase factor $(-1)^m$ in Eq. (10.27) for $m > 0$.

10.3 Radial Equation

We next come back to the radial Schrödinger equation, Eq. (10.11), substituting $R(r) = \frac{\tilde{R}(r)}{r}$ and remembering that $\lambda = l(l+1)$, we get

$$\Big[-\frac{\hbar^2}{2\mu}\frac{d^2}{dr^2} + V(r) + \frac{l(l+1)\hbar^2}{2\mu r^2}\Big]\tilde{R}(r) = E\ \tilde{R}(r) \tag{10.30}$$

This is an effective one-dimensional Schrödinger equation in r. However the interval is $0 \leq r < \infty$ only. Another important requirement is $\tilde{R}(0) = 0$, since $\tilde{R}(r) = r\,R(r)$ and $R(r)$ must be finite. Physical interpretation of Eq. (10.30) is the following. It shows that the particle of effective mass μ moves in the one-dimensional r-space with an effective potential $V_{eff}(r) = V(r) + \frac{l(l+1)\hbar^2}{2\mu r^2}$. The second term is the centrifugal repulsion. To visualize this, we can imagine an infinitely long and infinitesimally thin tube, along which the particle moves (for r-motion) with one end fixed at the origin. As the particle moves along this tube in an effective one-dimensional motion, it feels the centrifugal repulsion as the tube rotates about the origin giving rise to Eq. (10.30).

Although no minimum well depth is needed for the existence of a bound state for one-dimensional potentials (see Chap. 8), a minimum well depth is necessary in three-dimensional problems even when there is no centrifugal repulsion, i.e. for $l = 0$. This is because the r-space is restricted to $0 \leq r < \infty$ together with an additional requirement $\tilde{R}(0) = 0$. Thus even the ground state of Eq. (10.30) appears to have a node at $r = 0$ in the *fictitious one-dimensional problem in the interval* $-\infty < r < \infty$ *and replacing the potential by* $V(|r|)$. Then this state becomes an effective first excited state in this fictitious case, which needs a minimum well depth. However the physical interval is $0 \leq r < \infty$ and $r = 0$ is *not considered as a real node*.

Connection between physics and mathematics

1. The associated Legendre equation has regular singularities at $\theta = 0$ and π (corresponding to $w = -1$ and 1). According to Chap. 7, Sect. 7.1.4, the interval $[0, \pi]$ for θ or $[-1, 1]$ for w and weight function (equal to $\sin\theta$ for θ, or 1 for w) make the corresponding operator Hermitian and its eigen value real. These conditions coincide with the physical interval $[0, \pi]$ and weight function in Eq. (10.26).
2. Since $w = \pm 1$ are regular singular points, a series solution about $w = 0$ diverges at $w = \pm 1$, if the series is allowed to be an infinite series. Since quantum mechanical wave function must be finite everywhere in the physical interval,

this series must be terminated, which restricts the eigen value λ to $l(l+1)$ with $l = 0, 1, 2, \cdots$, which is experimentally verified. Once again, we see a perfect harmony between physics and mathematics.
3. The second solution of the associated Legendre equation (denoted by $Q_l^m(w)$) diverges at $w = \pm 1$, and hence it is rejected. This is just as it should be, as otherwise we would have more solutions than needed, hence non-unique.
4. Although there is no singularity in the eigen value equation for $\hat{L}_z$ [see Eq. (10.14)], the requirement of uniqueness of the wave function after a complete cycle selects the eigen value m to be an integer (positive, negative or zero), which is verified experimentally.
5. Equation (10.14) is a first order differential equation and has only one solution. This does not lead to any possibility of non-uniqueness, in agreement with the requirement of quantum mechanics.

10.4 Problems

1. If an operator $\hat{A}$ commutes with $\hat{L}_x$ and $\hat{L}_y$, show that $\hat{A}$ commutes with $\hat{L}_z$ and $\hat{L}^2$ also.
2. Show that the only state for which all three components of the orbital angular momentum ($\vec{L}$) can be specified, must be a state with $L = 0$.
3. Defining the operators

$$\hat{L}_+ = \hat{L}_x + i\hat{L}_y, \quad \hat{L}_- = \hat{L}_x - i\hat{L}_y = (\hat{L}_+)^\dagger,$$

calculate

$$[\hat{L}_+, \hat{L}_-], \quad [\hat{L}_+, \hat{L}_x], \quad [\hat{L}_+, \hat{L}_z], \quad [\hat{L}_+, \hat{L}_y^2].$$

4. Consider a quantum mechanical particle, restricted in the conical space defined by the cone about the positive z-axis with polar angle θ_0. Write down the Schrödinger equation and discuss how it can be solved, subject to the boundary conditions. Comment on the footnote following Eq. (7.26).

References

Chattopadhyay, P.K.: Mathematical Physics. New Age International (P) Ltd., New Delhi (2006)
Merzbacher, E.: Quantum Mechanics. John Wiley & Sons Inc., New York (1965)
Roy, R.R., Nigam, B.P.: Nuclear Physics: Theory and Experiment. John Wiley & Sons, New York (1967)
Schiff, L.I.: Quantum Mechanics, 2nd edn. McGraw-Hill Book Company Inc. (1955); Reprinted as International Student Edition, Kogakusha Co. Ltd. Tokyo

Chapter 11
Hydrogen-type Atoms: Two Bodies with Mutual Force



Abstract It is shown that the two-body motion with mutual interaction reduces to a motion in relative variable. This is solved for the simple hydrogen atom, ignoring spins.

Keywords Reduction of mutually interacting two-body system · Eigen values and eigen functions of H-type atoms

Hydrogen atom consists of an electron moving around a proton (nucleus). In the first simple treatment, we take the force acting between the particles to be the attractive Coulomb (electrostatic) force. Since Coulomb attraction is by far the most dominant force for the hydrogen (H) atom, this simple model is a very good approximation for the real H-atom and historically it provided one of the first concrete experimental verification of quantum mechanics in the first quarter of the twentieth century. More realistically, one has to take the spins of the electron and proton, relativistic effects, etc., into account. The predictions arising from inclusion of all these have been verified with extremely high precision in sophisticated experiments, over the years. However these effects are small and can be treated as perturbations.

The simple model is a two-body problem with mutual force – the electron (e) and proton (p) interacting via an electrostatic (e.s.) attractive force, which is directed along a line joining the two particles. Intuition dictates that properties of the H-atom should not depend on the individual positions of the particles, but on the relative separation vector. In the following we see that, in the absence of external forces, the center-of-mass (c.m.) motion separates and the problem reduces to a one-body problem.

11.1 Two Mutually Interacting Particles: Reduction to One-Body Schrödinger Equation

We consider the electron and proton as particles numbered 1 and 2, with masses m_1 and m_2 at position vectors $\vec{r}_1$ and $\vec{r}_2$ (relative to an arbitrary origin) respectively and interacting via a general potential $V(\vec{r}_1, \vec{r}_2)$. The two-body Schrödinger equation is

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5_11

$$\Big[-\frac{\hbar^2}{2m_1}\nabla^2_{\vec{r}_1}-\frac{\hbar^2}{2m_2}\nabla^2_{\vec{r}_2}+V(\vec{r}_1,\vec{r}_2)\Big]\Psi(\vec{r}_1,\vec{r}_2)=E_{\mathrm{T}}\Psi(\vec{r}_1,\vec{r}_2), \tag{11.1}$$

where E_{T} ia the *total* energy of the system and $\nabla^2_{\vec{r}_1}$, $\nabla^2_{\vec{r}_2}$ are the Laplacians corresponding to vectors $\vec{r}_1, \vec{r}_2$ respectively. Now, if the potential depends on the *relative separation vector only*, i.e. $V(\vec{r}_1,\vec{r}_2)=V(\vec{r}_1-\vec{r}_2)$, we can see that the motion of the system separates into the motion of the c.m. and a relative motion (Schiff 1955). Introduce a c.m. vector ($\vec{\mathcal{R}}$) and a relative vector ($\vec{r}$) through

$$\begin{aligned}&\vec{\mathcal{R}}=\frac{m_1\vec{r}_1+m_2\vec{r}_2}{M}, \qquad M=m_1+m_2,\\ &\vec{r}=\vec{r}_1-\vec{r}_2.\end{aligned} \tag{11.2}$$

Here M is the total mass of the system. Then it can be verified (Schiff 1955) that

$$\begin{aligned}&\frac{1}{m_1}\nabla^2_{\vec{r}_1}+\frac{1}{m_2}\nabla^2_{\vec{r}_2}=\frac{1}{M}\nabla^2_{\vec{\mathcal{R}}}+\frac{1}{\mu}\nabla^2_{\vec{r}},\\ \text{where}\quad &\frac{1}{\mu}=\frac{1}{m_1}+\frac{1}{m_2},\end{aligned} \tag{11.3}$$

and $\nabla^2_{\vec{\mathcal{R}}}$ and $\nabla^2_{\vec{r}}$ are the Laplacians corresponding to $\vec{\mathcal{R}}$ and $\vec{r}$ respectively. The quantity μ is called the reduced mass of the system. Substitution of Eq. (11.3) in Eq. (11.1) with $V(\vec{r}_1,\vec{r}_2)=V(\vec{r})$ gives

$$\Big[-\frac{\hbar^2}{2M}\nabla^2_{\vec{\mathcal{R}}}-\frac{\hbar^2}{2\mu}\nabla^2_{\vec{r}}+V(\vec{r})\Big]\Psi(\vec{\mathcal{R}},\vec{r})=E_{\mathrm{T}}\Psi(\vec{\mathcal{R}},\vec{r}).$$

Writing $\Psi(\vec{\mathcal{R}},\vec{r})=U(\vec{\mathcal{R}})u(\vec{r})$ and dividing by $\Psi=Uu$, we immediately see that the variables $\vec{\mathcal{R}}$ and $\vec{r}$ are separated

$$\begin{aligned}&\Big[-\frac{\hbar^2}{2\mu}\nabla^2_{\vec{r}}+V(\vec{r})\Big]u(\vec{r})=Eu(\vec{r})\\ &\Big[-\frac{\hbar^2}{2M}\nabla^2_{\vec{\mathcal{R}}}\Big]U(\vec{\mathcal{R}})=E_{\mathrm{cm}}U(\vec{\mathcal{R}})\\ \text{where}\quad &E_{\mathrm{T}}=E+E_{\mathrm{cm}}.\end{aligned} \tag{11.4}$$

Here E is the new separation constant. Left side of first of Eq. (11.4) can be interpreted to be the Hamiltonian and E to be the energy of the relative motion. The next equation shows that the center of mass moves as a *free particle* (since there is no potential) of mass M and energy E_{cm} ($=E_{\mathrm{T}}-E$). This is exactly what we expect, since there is no external force on the e-p system. Our interest is on the e-p system itself and not on the total kinetic energy (E_{T}) of the system as a whole, which includes *the motion of the center of mass with energy* E_{cm}. Hence we need to solve the first of Eq. (11.4) only. Thus the relevant motion reduces to an *effective one-body-motion* of

a fictitious particle of mass μ located at $\vec{r}$, moving in the potential field $V(\vec{r})$. As intuition demands, this is always true whenever a number of particles $N \geq 2$ interact among themselves via *mutual forces only* and there is no external force acting on the system. Then the motion of the center of mass of the system separates and it can be ignored. Properties of this system depend only on the relative motion of an effective $(N-1)$-body fictitious system.

For the H-atom, since the proton is approximately 1837 times heavier than the electron, we see from Eqs. (11.2) and (11.3) that the c.m. is almost coincident with the position of proton and the reduced mass is approximately equal to the electron mass.

11.2 Relative Motion of One-Electron H-Type Atoms

We will follow the general procedure outlined in Chap. 9, Sect. 9.3. The potential $V(\vec{r})$ of the H-atom arises from the Coulomb attraction of the electron with charge $-e$ and the proton with charge $+e$. We consider the more general case of one-electron H-type atoms: one-electron-atom with nuclear charge $+Ze$. For example, the Helium (He) atom contains two protons and two neutrons at its nucleus. Hence for the singly ionized He-atom $Z = 2$. Similarly $Z = 3$ for doubly ionized Lithium (Li) atom, etc. Then the e.s. potential is $V(\vec{r}) = -\frac{Ze^2}{r}$ and the relative Schrödinger equation is

$$\Big[-\frac{\hbar^2}{2\mu}\nabla_{\vec{r}}^2 - \frac{Ze^2}{r}\Big]u(\vec{r}) = Eu(\vec{r}) \tag{11.5}$$

The potential is spherically symmetric. The appropriate coordinate system is the spherical polar coordinates with the origin coinciding with the center of mass. According to our discussion in Chap. 10, the Hamiltonian commutes with the orbital angular momentum. Hence $u(\vec{r})$ is a simultaneous eigen function $Y_{lm}(\theta, \phi)$ of $\hat{L}^2$ and $\hat{L}_z$

$$u(\vec{r}) = R(r)Y_{lm}(\theta, \phi) \equiv \frac{\tilde{R}(r)}{r}Y_{lm}(\theta, \phi).$$

We include an additional factor $\frac{1}{r}$ in the radial wave function $R(r)$, to remove the first derivative in r from the radial Schrödinger equation (10.11), leading to Eq. (10.30). With $V(r) = -\frac{Ze^2}{r}$ this gives

$$\Big[-\frac{\hbar^2}{2\mu}\frac{d^2}{dr^2} - \frac{Ze^2}{r} + \frac{l(l+1)\hbar^2}{2\mu r^2}\Big]\tilde{R}(r) = E\ \tilde{R}(r). \tag{11.6}$$

Following the general procedure outlined in Chap. 9, Sect. 3, we first put the 1-D Schrödinger equation in a dimensionless form by writing $\rho = \alpha r$, with α having dimension $[L^{-1}]$ and defining β as

$$\rho = \alpha r, \quad \text{and} \quad \beta = \frac{2\mu Z e^2}{\hbar^2 \alpha}.$$

We have from Eq. (11.6)

$$\Big[\frac{d^2}{\mathrm{d}\rho^2} + \frac{\beta}{\rho} - \frac{l(l+1)}{\rho^2}\Big]\tilde{R}(\rho) = -\frac{2\mu E}{\hbar^2\alpha^2}\tilde{R}(\rho).$$

Now we choose α such that

$$-\frac{2\mu E}{\hbar^2\alpha^2} = \frac{1}{4}, \quad \text{so that} \quad \alpha = +\sqrt{-\frac{8\mu E}{\hbar^2}}. \tag{11.7}$$

Note that for bound states $E < 0$ and α is real and positive (as ρ, as well as, r are positive, being measures of physical distances). Then the left side of the first relation above is a positive dimensionless constant. We choose it to be $\frac{1}{4}$, so that the resulting differential equation [see Eq. (11.11)] is in a standard form, *viz.* the associated Laguerre equation. Then β becomes

$$\beta = \sqrt{-\frac{Z^2 e^4 \mu}{2\hbar^2 E}}. \tag{11.8}$$

With these definitions, the Schrödinger equation becomes

$$\Big[\frac{d^2}{\mathrm{d}\rho^2} + \frac{\beta}{\rho} - \frac{l(l+1)}{\rho^2} - \frac{1}{4}\Big]\tilde{R}(\rho) = 0. \tag{11.9}$$

Next we extract the $\rho \to 0$ and $\rho \to \infty$ behaviors. In the limit $\rho \to 0$ the dominant form of Eq. (11.9) for $l \neq 0$ is

$$\Big[\frac{d^2}{\mathrm{d}\rho^2} - \frac{l(l+1)}{\rho^2}\Big]\tilde{R}(\rho) = 0,$$

whose general solutions are ρ^{l+1} and ρ^{-l}. The second solution must be discarded as it diverges as $\rho \to 0$. For $l = 0$, the treatment is complicated, since in this case we cannot disregard the $\frac{\beta}{\rho}$ term Merzbacher (1965). However we note that $\tilde{R}(\rho) \propto \rho$ is an approximate solution of Eq. (11.9) with $l = 0$, for small ρ, when β is not too large. Moreover it satisfies the boundary condition $\tilde{R}(0) = 0$. Hence we can take $\tilde{R}(\rho) \propto \rho^{l+1}$ in both cases. Next the $\rho \to \infty$ limit of Eq. (11.9) is

$$\Big[\frac{d^2}{\mathrm{d}\rho^2} - \frac{1}{4}\Big]\tilde{R}(\rho) = 0,$$

whose general solutions are $\mathrm{e}^{-\frac{\rho}{2}}$ and $\mathrm{e}^{\frac{\rho}{2}}$. The growing exponential has to be discarded as it is not finite for $\rho \to \infty$. Separating these two functions, we write

$$\tilde{R}(\rho) = \rho^{l+1} \mathrm{e}^{-\frac{\rho}{2}} v(\rho), \tag{11.10}$$

where $v(\rho)$ is a new unknown function. As in the general case, separation of the two asymptotic behaviors will cancel corresponding terms from the differential equation for $v(\rho)$. Substituting Eq. (11.10) in Eq. (11.9), we get

$$\frac{d^2 v(\rho)}{\mathrm{d}\rho^2} + \Big(2\frac{(l+1)}{\rho} - 1\Big)\frac{\mathrm{d}v(\rho)}{\mathrm{d}\rho} + \Big(\frac{\beta}{\rho} - \frac{l+1}{\rho}\Big)v(\rho) = 0.$$

Multiplying by ρ

$$\rho\frac{d^2 v(\rho)}{\mathrm{d}\rho^2} + (2l + 2 - \rho)\frac{\mathrm{d}v(\rho)}{\mathrm{d}\rho} + (\beta - l - 1)v(\rho) = 0. \tag{11.11}$$

We can recognize this equation to be the associated Laguerre differential equation (7.52), whose solutions are the associated Laguerre functions. From the standard result [see Ref. Schiff (1955)] we know that unless $\beta - l - 1$ is a non-negative integer, the total wave function $\tilde{R}(\rho) = \rho^{l+1}\mathrm{e}^{-\frac{\rho}{2}} v(\rho)$ diverges as $\rho \to \infty$. Without using this knowledge, we can see it by adopting the general procedure outlined in Chap. 9, Sect. 3, which follows.

We try a series solution

$$v(\rho) = \sum_{i=0}^{\infty} a_i \rho^{k+i} \qquad (a_0 \neq 0).$$

Substituting in Eq. (11.11),

$$\sum_{i=0}^{\infty} a_i (k+i)(k+i-1)\rho^{k+i-1} + (2l+2)\sum_{i=0}^{\infty} a_i (k+i)\rho^{k+i-1}$$
$$- \sum_{i=0}^{\infty} a_i (k+i)\rho^{k+i} + (\beta - l - 1)\sum_{i=0}^{\infty} a_i \rho^{k+i} = 0.$$

The indicial equation is obtained by setting the coefficient of the lowest power of ρ, *viz.* ρ^{k-1} to zero

$$a_0 k(k-1) + (2l+2)a_0 k = 0.$$

Since $a_0 \neq 0$, we have

$$k(k + 2l + 1) = 0.$$

Since k cannot be negative (for which the wave function diverges at $\rho = 0$), we must have

$$k = 0.$$

The recurrence relation is obtained by setting the coefficient of $\rho^{k+i} = \rho^i$ to zero. This gives

$$\frac{a_{i+1}}{a_i} = \frac{i + l + 1 - \beta}{(i+1)(i+2l+2)}. \tag{11.12}$$

Thus starting with an arbitrary a_0, we can get an infinite series solution for $v(\rho)$. We next study the behavior of $v(\rho)$ if this series does not terminate. Then for very large i, the ratio of successive terms of the series becomes

$$\lim_{i\to\infty} \frac{t_{i+1}}{t_i} = \lim_{i\to\infty} \frac{a_{i+1}}{a_i}\rho \to \frac{\rho}{i},$$

using Eq. (11.12). We compare this ratio of successive terms of the tail part of this series with that of the series expansion of $\mathrm{e}^\rho = \sum_{i=0}^{\infty} \frac{\rho^i}{i!} \equiv \sum_{i=0}^{\infty} t_i'$

$$\lim_{i\to\infty} \frac{t_{i+1}'}{t_i'} = \lim_{i\to\infty} \frac{\rho}{i+1} \to \frac{\rho}{i}.$$

Comparing the tail parts of these two series, we find that the series for $v(\rho)$ behaves as e^ρ, if it is allowed to continue to infinity. Then contribution to the complete radial wave function coming from the tail part of the infinite series for $v(\rho)$ behaves as

$$\tilde{R}(\rho) \to \rho^{l+1}\mathrm{e}^{-\rho/2}\mathrm{e}^\rho = \rho^{l+1}\mathrm{e}^{\rho/2},$$

which diverges for $\rho \to \infty$ [this is a consequence of the singularity of Eq. (11.11) at $\rho = \infty$]. Thus we cannot allow the series to be an infinite series. If the series is finite, Eq. (11.10) shows that $\tilde{R}(\rho)$ vanishes in the limit $\rho \to \infty$. The series will terminate at $i = N$, with N an integer ≥ 0, if $a_{N+1} = 0$, while $a_N \neq 0$. We see from Eq. (11.12) that this can be achieved by choosing β such that

$$\beta = N + l + 1 \equiv n. \tag{11.13}$$

Since N and l are both integers ≥ 0, n is an integer ≥ 1. The quantum number n is called the *principal quantum number*. Substituting in Eq. (11.8), we have

$$E \equiv E_n = -\frac{Z^2 e^4 \mu}{2\hbar^2 n^2}, \qquad (n = 1, 2, 3, \cdots). \tag{11.14}$$

This is the energy quantization relation. We see that this results from the imposition of the boundary condition on the wave function at infinity. Thus the system cannot have any energy, but must have quantized energy eigen values given by Eq. (11.14). This relation gives rise to the H-atom spectrum, experimental verification of which was one of the original triumphs of quantum mechanics. Note that the old quantum theory also results in the same energy eigen values. Also note that Eq. (11.14) is a very good approximation, as the additional terms in the Hamiltonian neglected in

this simple treatment are quite small.

Eigen function:
To get the wave function, we go back to Eq. (11.11), with $\beta = n$

$$\rho \frac{d^2 v(\rho)}{\mathrm{d}\rho^2} + (2l + 2 - \rho)\frac{\mathrm{d}v(\rho)}{\mathrm{d}\rho} + (n - l - 1)v(\rho) = 0. \tag{11.15}$$

This can be recognized to be the standard associated Laguerre differential equation [see Eq. (7.52)]

$$\rho \frac{d^2 L_q^p(\rho)}{\mathrm{d}\rho^2} + (p + 1 - \rho)\frac{\mathrm{d}L_q^p(\rho)}{\mathrm{d}\rho} + (q - p)L_q^p(\rho) = 0, \tag{11.16}$$

where p and q are two constants. The definitions of associated Laguerre equation and its solutions are not unique. We follow the definition given by Schiff except for an unimportant overall sign of its solution, [see Ref. Schiff (1955)]. Comparing Eq. (11.16) with Eq. (11.15) we find that Eq. (11.15) is indeed the associated Laguerre differential equation with $p = 2l + 1$ and $q - p = n - l - 1$, i.e. $q = n + l$. Hence

$$v(\rho) = L_{n+l}^{2l+1}(\rho).$$

When q and p are integers (as in the present case), the associated Laguerre function, $L_q^p(\rho)$, becomes a polynomial, as was obtained for the series solution, by the requirement that the series solution terminates. This led to Eq. (11.13). Introducing the quantum numbers (n, l) in Eq. (11.10), the normalized radial eigen function $R_{nl}(r) = \frac{\tilde{R}_{nl}(r)}{r}$ becomes [see Ref. Schiff (1955)]

$$\begin{aligned} R_{nl}(r) &= \left[\left(\frac{2Z}{n a_{\text{Bohr}}}\right)^3 \frac{(n - l - 1)!}{2n\{(n + l)!\}^3}\right]^{\frac{1}{2}} \mathrm{e}^{-\frac{1}{2}\rho} \rho^l L_{n+l}^{2l+1}(\rho), \\ \text{with } a_{\text{Bohr}} &= \frac{\hbar^2}{\mu e^2}, \quad \text{and } \rho = \alpha r = \left(\frac{2Z}{n a_{\text{Bohr}}}\right) r. \end{aligned} \tag{11.17}$$

Here a_{Bohr}, called the *Bohr radius*, is the radius of the first circular orbit for H-atom $(Z = 1)$ in the old quantum theory. Hence the complete eigen function is

$$u_{nlm}(\vec{r}\,) = R_{nl}(r) Y_{lm}(\theta, \phi), \tag{11.18}$$

where $Y_{lm}(\theta, \phi)$ is the normalized spherical harmonics. The wave function is orthonormalized according to

$$\int u_{nlm}^*(\vec{r}\,)\, u_{n'l'm'}(\vec{r}\,)\, d^3 r = \delta_{nn'}\delta_{ll'}\delta_{mm'}. \tag{11.19}$$

Since under parity operation $(r, \theta, \phi) \rightarrow (r, \pi - \theta, \pi + \phi)$, the parity of $u_{nlm}(\vec{r})$ is given by that of $Y_{lm}(\theta, \phi)$, Eq. (10.29). Hence

$$u_{nlm}(\vec{r}) \underset{parity}{\longrightarrow} u_{nlm}(-\vec{r}) = (-1)^l u_{nlm}(\vec{r}). \tag{11.20}$$

Connection between physics and mathematics Following observations show the intimate connection between physics of this problem and the inherent mathematics:

1. The domain of regularity (without any intervening singularity) in the r space of the Schrödinger equation [as also that of the associated Laguerre differential equation (7.52)] is $0 \leq r < \infty$, with a regular singularity at $r = 0$ and an irregular singularity at $r = \infty$. The *physical domain* agrees exactly with this mathematical domain of regularity, as discussed after Eq. (7.26).
2. The eigen function $u_{nlm}(\vec{r})$ has a factor $e^{-\frac{1}{2}\alpha r}$. Now all measurable physical quantities are integrals involving bi-linear combination of $u^*_{nlm}(\vec{r})$ and $u_{n'l'm'}(\vec{r})$. Thus the radial integral has a factor $e^{-\alpha r}$. This agrees with the weight function $e^{-\alpha r}$ of the associated Laguerre differential equation [see Chap. 7].
3. According to Chap. 7, Sect. 7.1.4, the interval and the weight function of the associated Laguerre differential equation makes the corresponding differential operator (hence also the Hamiltonian) Hermitian, which is a basic requirement according to the fundamental postulates, Chap. 3, Sect. 3.2.
4. Physics requires the wave function to be finite everywhere within the domain. The fact that $r = \infty$ is an irregular singular point demands that the series solution of the associated Laguerre differential equation be terminated, which in turn gives rise to energy eigen values. Thus physical requirement and mathematical properties are closely related.

We see that the physical interval $0 \leq r < \infty$, the fact that $r = \infty$ is an irregular singular point and that the eigen function has the factor $e^{-\frac{1}{2}\alpha r}$ are just right from the mathematical point of view.

11.3 Problems

1. Calculate the probability of finding the electron within a thin shell of thickness dr, and having radius r for the states $(n = 1, l = 0)$, $(n = 2, l = 0)$ and $(n = 2, l = 1)$. Use the series expansion of L^{2l+1}_{n+l} given in Ref. Schiff (1955), with overall sign changed. Plot these as function of r.
2. Calculate expectation values of r and r^2 for the ground and first excited states of the H-atom.
3. Calculate the expectation value of $\frac{1}{r}$ for the n-th state of H-type atoms. From this calculate the expectation value of the potential energy. Use this to calculate the expectation value of the kinetic energy.

References

Merzbacher, E.: Quantum Mechanics. John Wiley & Sons Inc., New York (1965)
Schiff, L.I.: Quantum Mechanics, 2nd edn. McGraw-Hill Book Company Inc. (1955); reprinted as International Student Edition, Kogakusha Co. Ltd. Tokyo

Chapter 12
Particle in a 3-D Well

Abstract Specific examples of three-dimensional potentials have been discussed. They include spherically symmetric hole with rigid and permeable walls, a cylindrical hole with rigid walls and the three-dimensional spherically symmetric harmonic oscillator. We stress that the choice of a coordinate system consistent with the symmetry of the system simplifies the problem.

Keywords Spherically symmetric holes · With rigid and permeable walls · Cylindrical hole · Three-dimensional harmonic oscillator

In this chapter, we will consider a free quantum particle in a three-dimensional hole. Within the hole the particle is *free*, i.e. it is not acted upon by any force and the potential is a constant inside the hole. Since a constant shift in potential is equivalent to a shift in energy by the same constant amount (which is unimportant in non-relativistic motion), we take the potential within the hole to be zero. The walls of the hole may be rigid or soft. The rigid wall will not permit the particle (classical or quantum) to cross the wall. Hence the potential jumps to infinity from the hole to the wall. The potential at a soft wall increases discontinuously to a positive constant, say V_0. A classical particle with energy $E < V_0$ cannot penetrate the wall. But a quantum particle with any energy has a probability to be within this wall. In the following sections, we treat several such cases. The analytical treatment is simplified by the choice of the origin and the coordinate system according to the symmetry of the potential.

12.1 Spherically Symmetric Hole with Rigid Walls

As a simple example of a spherically symmetric three-dimensional hole, we consider a quantum particle of mass μ in a spherical hole of radius a with an infinitely rigid spherical wall at $r = a$. Because of the spherical symmetry, spherical polar coordinates are chosen with the center of the hole as the origin. Since the particle cannot penetrate the region $r > a$ and is free inside, we take the potential as

$$\begin{aligned} V(\vec{r}) &= 0 \quad \text{for} \quad r \le a, \\ &= \infty \quad \text{for} \quad r > a. \end{aligned} \tag{12.1}$$

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5_12

The Schrödinger equation for $r < a$ is

$$-\frac{\hbar^2}{2\mu}\nabla_{\vec{r}}^2\, u(\vec{r}\,) = E\, u(\vec{r}\,). \qquad (r < a)$$

For $r \geq a$, wave function vanishes, $u(\vec{r}\,) = 0$. The potential is spherically symmetric, and hence it commutes with $\hat{L}^2$ and $\hat{L}_z$. Hence the Hamiltonian also commutes with $\hat{L}^2$ and $\hat{L}_z$

$$[\hat{H}, \hat{L}^2] = 0, \qquad [\hat{H}, \hat{L}_z] = 0.$$

In this three-dimensional case, we have three quantum numbers: besides l and m, the radial quantum number is n and the eigen function is a simultaneous eigen function of $\hat{L}^2$ and $\hat{L}_z$, *viz.*, $Y_{lm}(\theta, \phi)$

$$u(\vec{r}\,) \equiv u_{nlm}(\vec{r}\,) = R_{nl}(r) Y_{lm}(\theta, \phi).$$

The radial Schrödinger equation for R, Eq. (10.11), for $r < a$, with $V(r) = 0$ and $\lambda = l(l+1)\hbar^2$ is

$$-\frac{\hbar^2}{2\mu}\Big[\frac{1}{r^2}\frac{d}{\mathrm{d}r}\Big(r^2\frac{d}{\mathrm{d}r}\Big)\Big]R(r) + \frac{l(l+1)\hbar^2}{2\mu r^2}R(r) = E\; R(r).$$

Next we put this equation in a dimensionless form by introducing a dimensionless variable $\rho = \alpha r$, with α having the dimension of $[L^{-1}]$. Multiplying above equation by $-\frac{2\mu}{\hbar^2\alpha^2}$ and setting

$$\frac{2\mu E}{\alpha^2\hbar^2} = 1 \;\text{ (dimensionless)}, \quad \text{hence } \alpha = \sqrt{\frac{2\mu E}{\hbar^2}},$$

we have for $R \equiv R(\rho)$

$$\frac{d^2R}{\mathrm{d}\rho^2} + \frac{2}{\rho}\frac{\mathrm{d}r}{\mathrm{d}\rho} + \Big[1 - \frac{l(l+1)}{\rho^2}\Big]R = 0. \qquad (12.2)$$

Multiplying by ρ^2, we recognize this as the spherical Bessel equation (7.69) and the general solutions are spherical Bessel ($j_l(\rho)$) and spherical Neumann ($n_l(\rho)$) functions (see Chap. 7, Sect. 7.2)

$$R(\rho) \equiv R_{nl}(\rho) = A_{nl}\, j_l(\rho) + B_{nl}\, n_l(\rho),$$

where A_{nl} and B_{nl} are two arbitrary constants to be determined from the boundary conditions. The quantum number l corresponds to a specific orbital angular momentum. We have introduced the quantum number n in anticipation of radial quantization. From the properties of spherical Bessel and spherical Neumann functions (Chap. 7, Sect. 7.2) we see that $n_l(\rho)$ diverges as $\rho \to 0$, while $j_l(\rho)$ is finite everywhere. Hence in order that the $R(\rho)$ is finite, we have to set $B_{nl} = 0$

$$R(\rho) \equiv R_{nl}(\rho) = A_{nl}\, j_l(\rho). \tag{12.3}$$

The next boundary condition is that the radial function must vanish at $r = a$

$$R(\rho)|_{r=a} = A_{nl}\, j_l\left(\sqrt{\frac{2\mu E_{nl}}{\hbar^2}}a\right) = 0. \tag{12.4}$$

From a plot of the functions, we see that $j_l(\rho)$ for a given l oscillates with decreasing amplitude as ρ increases [see Refs. Arfken (1966), Abramowitz and Stegun (1972)]. There are infinitely many zeros: $j_l(z_{ln}) = 0$ with $n = 1, 2, 3, \cdots$. Hence Eq. (12.4) will be satisfied if

$$\sqrt{\frac{2\mu E_{nl}}{\hbar^2}}a = z_{ln} \quad \text{or} \quad E_{nl} = \frac{\hbar^2 z_{ln}^2}{2\mu a^2}, \quad n = 1, 2, 3, \cdots. \tag{12.5}$$

[Numerical values of the zeros of $j_l(\rho)$ are tabulated in Ref. Abramowitz and Stegun (1972)]. Thus energy for a given l is quantized. Substituting E_{nl} from Eq. (12.5) in the expression for α, we have for specified n, l

$$\alpha \equiv \alpha_{ln} = \frac{z_{ln}}{a},$$

and the corresponding eigen function

$$u_{nlm}(\vec{r}\,) = A_{nl}\, j_l\left(\frac{z_{ln}}{a}r\right) Y_{lm}(\theta, \phi). \tag{12.6}$$

The constant A_{nl} can be found from the standard normalization condition.

For the special case of $l = 0$, we have Arfken (1966)

$$j_0(\rho) = \frac{\sin\rho}{\rho}.$$

Hence solution of Eq. (12.4) for $l = 0$ is $z_{0n} = n\pi$, $n = 1, 2, \cdots$, and we have energy eigen values for $l = 0$

$$E_{n0} = \frac{\hbar^2 n^2 \pi^2}{2\mu a^2}, \quad n = 1, 2, 3, \cdots \quad \text{for} \quad l = 0.$$

This can also be obtained by a simple solution of Eq. (10.30) with $V = 0$ and $l = 0$.

For a given l, it is apparent that energy increases with n. Since the centrifugal term is repulsive, energy increases with l also. For the lowest set of eigen values, we have to take $l = 0$. The minimum energy of the particle in this hole (with $l = 0$, $n = 1$) is thus $E_{10} = \frac{\hbar^2\pi^2}{2\mu a^2}$. The uncertainty principle requires a zero point motion, leading to this minimum energy.

Mathematical note

The nth zero of $j_l(\rho)$ and $n_l(\rho)$ are counted from the first zero after $\rho = 0$. Since $j_l(0) = 0$ for all $l \neq 0$ and $j_0(0) = 1$, normally one would expect $\rho = 0$ as the first zero of $j_l(\rho)$ for $l \neq 0$, i.e., $z_{l0} = 0$. But then the eigen function, Eq. (12.6) would vanish identically for $n = 0$. This is a trivial, non-physical solution. Hence the zeros are counted successively as ρ increases for $\rho > 0$.

12.2 Spherically Symmetric Hole with Permeable Walls

We next consider the above case with soft (permeable) walls, so that the potential is zero inside the spherical hole of radius a, but V_0 outside (V_0 positive and $E < V_0$ for a bound state)

$$\begin{aligned} V(\vec{r}\,) &= 0 \text{ for } r < a \text{ (Region I)} \\ &= V_0 \text{ for } r \geq a \text{ (Region II)}. \end{aligned}$$

For $r < a$, the solution is exactly as in the previous section and is given by Eq. (12.3)

$$R_{nl}^{(\mathrm{I})}(r) = A_{nl}\, j_l(\alpha r) \quad \text{with } \alpha = \sqrt{\frac{2\mu E}{\hbar^2}}, \qquad \text{for } r < a. \tag{12.7}$$

However the wave function does not vanish at $r = a$. The Schrödinger equation for $r \geq a$ is

$$-\frac{\hbar^2}{2\mu}\nabla_{\vec{r}}^2\, u^{(\mathrm{II})}(\vec{r}\,) + V_0\, u^{(\mathrm{II})}(\vec{r}\,) = E\, u^{(\mathrm{II})}(\vec{r}\,). \qquad (r \geq a)$$

The radial Schrödinger equation for $R^{(\mathrm{II})}$, Eq. (10.11), for $r \geq a$, with $V(r) = V_0$ and $\lambda = l(l+1)\hbar^2$ is

$$-\frac{\hbar^2}{2\mu}\left[\frac{1}{r^2}\frac{d}{\mathrm{d}r}\left(r^2\frac{d}{\mathrm{d}r}\right)\right]R^{(\mathrm{II})}(r) + \left[V_0 + \frac{l(l+1)\hbar^2}{2\mu r^2}\right]R^{(\mathrm{II})}(r) = E\, R^{(\mathrm{II})}(r).$$

For a bound state, we must have $E < V_0$, hence

$$\frac{d^2}{\mathrm{d}r^2}R^{(\mathrm{II})}(r) + \frac{2}{r}\frac{d}{\mathrm{d}r}R^{(\mathrm{II})}(r) - \left[\beta^2 + \frac{l(l+1)}{r^2}\right]R^{(\mathrm{II})}(r) = 0,$$

where $\beta = \sqrt{\frac{2\mu(V_0 - E)}{\hbar^2}}$ is a real positive constant, as $E < V_0$.

As before, we put this equation in a dimensionless form by introducing a dimensionless variable $\rho = \beta r$, with β having the dimension of $[L^{-1}]$. In terms of the new variable ρ, we have for $R^{(\mathrm{II})} \equiv R^{(\mathrm{II})}(\rho)$

$$\rho^2 \frac{d^2 R^{(\mathrm{II})}}{\mathrm{d}\rho^2} + 2\rho \frac{\mathrm{d}R^{(\mathrm{II})}}{\mathrm{d}\rho} - \big[\rho^2 + l(l+1)\big] R^{(\mathrm{II})} = 0. \tag{12.8}$$

We recognize this as the spherical modified Bessel equation (7.76) whose linearly independent standard solutions are spherical modified Bessel functions of first kind $[i_l(\rho)]$ and second kind $[k_l(\rho)]$. Hence a general solution is

$$R^{(\mathrm{II})}(\rho) \equiv R^{(\mathrm{II})}_{nl}(\rho) = C_{nl}\, i_l(\rho) + D_{nl}\, k_l(\rho), \qquad (\beta a \le \rho < \infty),$$

where C_{nl} and D_{nl} are two arbitrary constants to be determined from the boundary conditions. We have introduced the quantum number l for the specific orbital angular momentum, as also n in anticipation of radial quantization. Now for large ρ the function $i_l(\rho)$ diverges, while $k_l(\rho)$ is finite [see Ref. Arfken (1966)], hence we have $C_{nl} = 0$. In terms of r

$$R^{(\mathrm{II})}_{nl}(r) = D_{nl}\, k_l(\beta r) \quad \text{with } \beta = \sqrt{\frac{2\mu(V_0 - E)}{\hbar^2}} \quad (a \le r < \infty). \tag{12.9}$$

Next the continuity conditions at $r = a$ are

$$\begin{aligned} R^{(\mathrm{I})}_{nl}(r)\Big|_{r=a} &= R^{(\mathrm{II})}_{nl}(r)\Big|_{r=a} \\ \frac{dR^{(\mathrm{I})}_{nl}(r)}{dr}\Big|_{r=a} &= \frac{dR^{(\mathrm{II})}_{nl}(r)}{dr}\Big|_{r=a}. \end{aligned} \tag{12.10}$$

The arbitrary constants A_{nl} and D_{nl} can be eliminated by taking the ratio of these equation (the continuity of log-derivative of the wave function)

$$\Big(\frac{1}{R^{(\mathrm{I})}_{nl}(r)} \frac{dR^{(\mathrm{I})}_{nl}(r)}{dr}\Big)\Big|_{r=a} = \Big(\frac{1}{R^{(\mathrm{II})}_{nl}(r)} \frac{dR^{(\mathrm{II})}_{nl}(r)}{dr}\Big)\Big|_{r=a}.$$

Substituting from Eqs. (12.7) and (12.9)

$$\frac{\alpha\, j_l'(\alpha a)}{j_l(\alpha a)} = \frac{\beta\, k_l'(\beta a)}{k_l(\beta a)}$$

where $\alpha = \sqrt{\frac{2\mu E_{nl}}{\hbar^2}}$, and $\beta = \sqrt{\frac{2\mu(V_0 - E_{nl})}{\hbar^2}}$.

Here we have replaced E by its full expression E_{nl}. This is a transcendental equation and has to be solved numerically for E_{nl}. For a given l, the minimum energy is denoted by E_{1l}, the next higher one by E_{2l}, and so on.

For the eigen function, Eq. (12.10) can be used to express the constant D_{nl} in terms of A_{nl} for selected values of l and E_{nl}. Finally this overall normalization constant can be found by the normalization condition of the complete wave function, comprising those in the two regions. In general numerical methods are to be used for these calculations.

12.3 A Particle in a Cylindrical Hole with Rigid Walls

We next consider a particle having mass μ confined within a right circular cylindrical hole with rigid walls. For the cylindrical symmetry, we choose cylindrical coordinates, $\vec{r} \to (r, \phi, z)$, with z-axis along the symmetry axis of the cylinder of radius a and height H. The origin is at the bottom of the cylinder, $z = 0$. The potential within the hole is chosen as zero, while that on the walls and outside is ∞. Thus

$$\begin{aligned} V(\vec{r}) &= 0 \quad (0 \le r < a,\ 0 \le \phi < 2\pi,\ 0 < z < H) \\ &= \infty \quad \text{outside the cylinder.} \end{aligned} \tag{12.11}$$

The Schrödinger equation within the cylinder is

$$-\frac{\hbar^2}{2\mu}\nabla_{\vec{r}}^2\, u(\vec{r}) = E\, u(\vec{r}). \qquad \text{(within the cylinder).}$$

We should have $E > 0$ for bound states, since the potential within the cylinder is zero. On the surfaces of the cylinder and outside it, the wave function vanishes, $u(\vec{r}) = 0$. In cylindrical coordinates [see Ref. Arfken (1966) for $\nabla_{\vec{r}}^2$ in cylindrical coordinates] the Schrödinger equation becomes

$$-\frac{\hbar^2}{2\mu}\Big[\frac{1}{r}\frac{\partial}{\partial r}\Big(r\frac{\partial}{\partial r}\Big) + \frac{1}{r^2}\frac{\partial^2}{\partial\phi^2} + \frac{\partial^2}{\partial z^2}\Big]u(r,\phi,z) - E\,u(r,\phi,z) = 0 \quad \text{(within cylinder).} \tag{12.12}$$

Since the potential (12.11) is *not spherically symmetric* (it has θ-dependence in polar coordinates), it does not commute with $\hat{L}^2$. On the other hand, $V(\vec{r})$ has no ϕ dependence, hence it commutes with $\hat{L}_z$. Hence

$$[\hat{H}, \hat{L}_z] = 0, \quad \text{but} \quad [\hat{H}, \hat{L}^2] \neq 0.$$

Thus, although orbital angular momentum quantum number l is not good, the azimuthal quantum number m is good. We can write $u(r, \phi, z)$ as a simultaneous eigen function of $\hat{L}_z$, which is, $e^{im\phi}$, up to a normalization constant, where m is an integer (positive, negative or zero) representing the azimuthal quantum number

$$u(r, \phi, z) = U(r, z)e^{im\phi}.$$

This can be obtained independently from Eq. (12.12) by separation of ϕ, with continuity condition at $\phi = 0$ and 2π. Substituting in eq.(12.12) and multiplying by $-\frac{2\mu}{\hbar^2}$ we get

$$\Big[\frac{1}{r}\frac{\partial}{\partial r}\Big(r\frac{\partial}{\partial r}\Big) - \frac{m^2}{r^2} + \frac{\partial^2}{\partial z^2}\Big]U(r, z) + \alpha^2 U(r, z) = 0,$$

where $\alpha = \sqrt{\frac{2\mu E}{\hbar^2}}$ is a real positive constant, since $E > 0$. Clearly U is separable in the z-variable. Writing $U(r, z) = R(r)Z(z)$ we have

$$\frac{1}{R(r)}\Big[\frac{d^2}{\mathrm{d}r^2} + \frac{1}{r}\frac{d}{\mathrm{d}r} - \frac{m^2}{r^2}\Big]R(r) + \alpha^2 = -\frac{1}{Z(z)}\frac{d^2Z(z)}{dz^2} = \beta^2.$$

where β^2 is a constant, since left most side is a function of r only and the middle term is a function of z only. As $Z(z)$ must vanish on the top and bottom surfaces, the solution must be oscillatory in z and hence β^2 must be a positive constant, hence β real. We take β to be a real positive constant. A general solution of the z-equation is

$$Z(z) = A\sin\beta z + B\cos\beta z,$$

where A and B are arbitrary constants.

Boundary condition at $z = 0$ is : $Z(0) = 0. \Rightarrow B = 0.$

Boundary condition at $z = H$ is : $Z(H) = A\sin\beta H = 0. \Rightarrow \beta = \dfrac{n\pi}{H}$

with $n = 1, 2, \cdots$ ($n = 0$ gives a trivial solution, Z(z) = 0).

With these, the radial equation becomes, after multiplying by $r^2R(r)$

$$r^2\frac{d^2R}{\mathrm{d}r^2} + r\frac{\mathrm{d}R}{\mathrm{d}r} + \Big[(\alpha^2 - \beta^2)r^2 - m^2\Big]R = 0.$$

Introducing the dimensionless variable $\rho = \sqrt{\alpha^2 - \beta^2}\, r$

$$\rho^2\frac{d^2R}{\mathrm{d}\rho^2} + \rho\frac{\mathrm{d}R}{\mathrm{d}\rho} + \Big[\rho^2 - m^2\Big]R(\rho) = 0. \tag{12.13}$$

We recognize this as the Bessel equation, Eq. (7.56), of order m. The standard linearly independent solutions are $J_m(\rho)$ and $N_m(\rho)$. Since the latter diverges at $\rho = 0$ (see Chap. 7, Sect. 7.2), only acceptable solution is $J_m(\rho)$

$$R(r) = C\,J_m(\sqrt{\alpha^2 - \beta^2}\, r), \tag{12.14}$$

where C is a normalization constant. Next, the wave function must vanish on the curved surface of the cylinder, i.e. at $r = a$

$$R(r = a) = C J_m(\sqrt{\alpha^2 - \beta^2}\, a) = 0. \quad \text{Hence } \sqrt{\alpha^2 - \beta^2}\, a = \zeta_{mp},$$

where ζ_{mp} (which is a dimensionless number) is the p-th zero of $J_m(\rho)$, $p = 1, 2, \cdots$ [see Refs. Arfken (1966), Abramowitz and Stegun (1972)]. Substituting for α and β, with E replaced by $E_{p,m,n}$, we have

$$E_{p,m,n} = \frac{\hbar^2}{2\mu}\left[\left(\frac{\zeta_{mp}}{a}\right)^2 + \left(\frac{n\pi}{H}\right)^2\right], \tag{12.15}$$

with $p = 1, 2, \cdots$, $n = 1, 2, \cdots$ and m an integer. Since for integer m, the Bessel functions $J_{-m}(\rho)$ is proportional to $J_m(\rho)$ (Arfken 1966; Chattopadhyay 2006), we can restrict the quantum number m to non-negative integers, i.e. $m = 0, 1, 2, \cdots$. However as l is *not a good quantum number*, all values of $l \geq |m|$ contribute to a particular state with quantum numbers p, m, n.

For the ground state $E_{p,m,n}$ should be minimum. Clearly it increases with n and p. From a Table of zeros of integral order Bessel functions [see Ref. (Arfken 1966, Abramowitz and Stegun 1972)], we find that ζ_{mp} has a minimum value for $m = 0$ and $p = 1$. Then for the ground state the quantum numbers are $p = 1, m = 0, n = 1$. Since $\zeta_{m=0,p=1} = 2.4048$ (Arfken 1966) the ground state energy is

$$E_{p=1,m=0,n=1} = \frac{\hbar^2}{2\mu}\left[\left(\frac{2.4048}{a}\right)^2 + \left(\frac{\pi}{H}\right)^2\right].$$

Also note that the quantity within square brackets in Eq. (12.15) is α^2 and is clearly greater than β^2, so that $\sqrt{\alpha^2 - \beta^2}$ is real. Combining solutions for three separated variables, we have the complete eigen function

$$\begin{aligned} u_{p,m,n}(r, \phi, z) &= D J_m(\sqrt{\alpha^2 - \beta^2}\, r)\ \sin \beta z\ \mathrm{e}^{im\phi} \\ &= D\, J_m\left(\frac{\zeta_{mp}}{a}\, r\right)\ \sin\left(\frac{n\pi z}{H}\right)\ \mathrm{e}^{im\phi}, \end{aligned} \tag{12.16}$$

where D is the overall normalization constant, obtained from the standard normalization condition, such that eigen functions corresponding to different quantum numbers p, m, n are orthonormal. Note that volume element for the cylindrical coordinates is $r\mathrm{d}r\,\mathrm{d}\phi\,\mathrm{d}z$, while limits of r, ϕ and z integrations are from 0 to a, from 0 to 2π and from 0 to H respectively in this problem. The orthonormality relation of $\left\{J_m\left(\frac{\zeta_{mp}}{a}\, r\right)\right\}$, with respect to different p is given by Eq. (5.46) of Ref. Chattopadhyay (2006).

12.4 3-D Spherically Symmetric Harmonic Oscillator

Consider a particle of mass μ in a spherically symmetric harmonic oscillator represented by the potential $V(\vec{r}) = \frac{1}{2}Kr^2$, where K is the stiffness constant of the oscillator. The Schrödinger equation is

$$\Big[-\frac{\hbar^2}{2\mu}\nabla_{\vec{r}}^2 + \frac{1}{2}Kr^2 \Big] u(\vec{r}) = E\, u(\vec{r}). \tag{12.17}$$

This equation is easily seen to be separable in Cartesian coordinates, since

$$r^2 = x^2 + y^2 + z^2,$$

and

$$\nabla_{\vec{r}}^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial x^2}.$$

Each of the separated equations in x-, y- and z-variable is identical to the one-dimensional harmonic oscillator (Chap. 9, Sect. 9.4), with independent quantum numbers n_x, n_y and n_z respectively for x, y and z motions. The total energy is the sum of contributions from each motion, given by equations similar to Eq. (9.24)

$$\begin{aligned} &E_{n_x,n_y,n_z} = (n_x + n_y + n_z + \frac{3}{2})\hbar\omega_c \equiv (N + \frac{3}{2})\hbar\omega_c \\ &\text{where } \omega_c = \sqrt{\frac{K}{\mu}} \text{ and } N = n_x + n_y + n_z, \\ &\text{with each of } n_x, n_y, n_z = 0, 1, 2, \cdots . \end{aligned} \tag{12.18}$$

Each of the quantum numbers n_x, n_y and n_z can take non-negative integral values $0, 1, 2, \cdots$. Hence $N = 0, 1, 2, \cdots$ and the complete eigen function is a product of eigen functions for the x-, y- and z-motions

$$u_{n_x,n_y,n_z}(\vec{r}) = u_{n_x}(x)u_{n_y}(y)u_{n_z}(z), \tag{12.19}$$

where each of the single variable eigen functions are given by Eq. (9.27) together with Eq. (9.33), with appropriate changes.

Degeneracy: The eigen value given by Eq. (12.18) for a given N (other than 0) is degenerate. A given N has the same value for different combinations of n_x, n_y and n_z adding up to N. Thus n_x can take integer values from 0 to N, and for a given n_x, the quantity n_y can take integer values from 0 to $N - n_x$, and then n_z has a fixed value for given n_x and n_y. Hence for a given N the degeneracy is

$$D(N) = \sum_{n_x=0}^{N} \sum_{n_y=0}^{N-n_x} 1 = \frac{1}{2}(N+1)(N+2). \tag{12.20}$$

Solution in spherical polar coordinates Although the above treatment is correct, it does not show the angular momentum due to the three-dimensional motion. We see that the potential in Eq. (12.17) is spherically symmetric. Hence, according to Chap. 10, orbital angular momentum operator commutes with the Hamiltonian

$$[\hat{\vec{L}}, \hat{H}] = 0.$$

Hence orbital angular momentum (l) and its z-projection (m) are good quantum numbers. The eigen function of $\hat{H}$ is a simultaneous eigen function of $\hat{L}^2$ and $\hat{L}_z$, viz., spherical harmonics

$$u_{N,l,m}(\vec{r}) = \frac{\tilde{R}(r)}{r} Y_{lm}(\theta, \phi) \equiv \frac{\tilde{R}_{Nl}(r)}{r} Y_{lm}(\theta, \phi),$$

in which we have introduced a third quantum number N for the radial motion. Note that we need three quantum numbers for the three-dimensional motion. The radial equation is given by Eq. (10.30) [we drop subscripts of $\tilde{R}_{Nl}(r)$ for brevity]

$$\Big[-\frac{\hbar^2}{2\mu} \frac{d^2}{\mathrm{d}r^2} + \frac{1}{2} K r^2 + \frac{l(l+1)\hbar^2}{2\mu r^2} \Big] \tilde{R}(r) = E\ \tilde{R}(r).$$

Following the standard technique, we introduce the dimensionless variable $\rho = \alpha r$ and the equation becomes

$$\frac{d^2 R(\rho)}{\mathrm{d}\rho^2} + \Big[\frac{2\mu E}{\hbar^2 \alpha^2} - \frac{l(l+1)}{\rho^2} - \frac{\mu K}{\hbar^2 \alpha^4} \rho^2 \Big] \tilde{R}(\rho) = 0.$$

Choose α such that

$$\frac{\mu K}{\hbar^2 \alpha^4} = 1 \quad \Rightarrow \quad \alpha = \Big(\frac{\mu K}{\hbar^2} \Big)^{\frac{1}{4}} = \sqrt{\frac{\mu \omega_c}{\hbar}},$$

where $\omega_c = \sqrt{\frac{K}{\mu}}$ is the classical frequency of the oscillator. We can verify that α has the dimension $[L^{-1}]$, as required. Substitute

$$\lambda = \frac{2\mu E}{\hbar^2 \alpha^2} \quad \Rightarrow \quad \lambda = \frac{2E}{\hbar \omega_c}, \tag{12.21}$$

which is dimensionless. Then the dimensionless radial equation in terms of ρ is

$$\frac{d^2 \tilde{R}(\rho)}{\mathrm{d}\rho^2} + \Big[\lambda - \frac{l(l+1)}{\rho^2} - \rho^2 \Big] \tilde{R}(\rho) = 0. \tag{12.22}$$

Next we investigate the dominant behavior for large r (large ρ). As in one-dimensional case, the dominant asymptotic part of this equation becomes

$$\frac{d^2R(\rho)}{\mathrm{d}\rho^2} - \rho^2 R(\rho) = 0,$$

whose acceptable dominant solution is $\mathrm{e}^{-\frac{1}{2}\rho^2}$. Hence write

$$\tilde{R}(\rho) = v(\rho)\mathrm{e}^{-\frac{1}{2}\rho^2}, \tag{12.23}$$

and substitute in Eq. (12.22) resulting in (canceling out a common factor $\mathrm{e}^{-\frac{1}{2}\rho^2}$)

$$\frac{d^2v}{\mathrm{d}\rho^2} - 2\rho\frac{dv}{\mathrm{d}\rho} + \left[(\lambda - 1) - \frac{l(l+1)}{\rho^2}\right]v = 0.$$

Next to extract the $\rho \to 0$ behavior, we substitute $v(\rho) = \rho^\beta \eta(\rho)$ (with $\beta \geq 0$ for the solution to be finite at $\rho = 0$) and find that the dominant $\rho \to 0$ term gives

$$\beta = l + 1 \quad \text{or} \quad -l.$$

Since β must be ≥ 0, we choose $\beta = l + 1$. Hence

$$v(\rho) = \rho^{l+1}\eta(\rho) \quad \text{and} \quad \tilde{R}(\rho) = \rho^{l+1}\mathrm{e}^{-\frac{1}{2}\rho^2}\eta(\rho). \tag{12.24}$$

The equation satisfied by η is

$$\rho\frac{d^2\eta}{\mathrm{d}\rho^2} + 2\Big[(l+1) - \rho^2\Big]\frac{\mathrm{d}\eta}{\mathrm{d}\rho} + \Big[(\lambda - 3) - 2l\Big]\rho\eta = 0.$$

Comparison with associated Laguerre equation (7.52) and noticing that ρ^2 appears in the second term suggest we transform the variable to $\zeta = \rho^2$, with $\eta(\rho) = \nu(\zeta)$. Substituting these the equation for ν becomes

$$\zeta\frac{d^2\nu}{\mathrm{d}\zeta^2} + \left[(l + \frac{1}{2}) + 1 - \zeta\right]\frac{\mathrm{d}\nu}{\mathrm{d}\zeta} + \left[\frac{\lambda}{4} - \frac{(2l+3)}{4}\right]\nu(\zeta) = 0.$$

Comparing with Eq. (7.52), we notice that the above equation is the associated Laguerre equation with (P, Q replaced by p, q)

$$p = l + \frac{1}{2} \quad \text{and} \quad q - p = \frac{\lambda}{4} - \frac{(2l+3)}{4} \quad \Rightarrow \quad q = \frac{\lambda}{4} + \frac{l}{2} - \frac{1}{4}.$$

Its solution is the associated Laguerre function (see Chap. 7, Sect. 7.2)

$$v(\zeta) = L_q^p(\zeta) = L^{l+\frac{1}{2}}_{\frac{\lambda}{4}+\frac{l}{2}-\frac{1}{4}}(\zeta).$$

From a series solution for $v(\rho^2)$ we see (as in the one-dimensional case) that the solution (12.24) will diverge for $\rho \to \infty$, if the series is allowed to be an infinite series. To make the solution finite as $\rho \to \infty$, we must terminate the series. From the discussion after Eq. (7.52), we see that $L^{l+\frac{1}{2}}_{\frac{\lambda}{4}+\frac{l}{2}-\frac{1}{4}}(\zeta)$ will be a polynomial in $\zeta = \rho^2$ of degree $q - p = \frac{\lambda}{4} - \frac{l}{2} - \frac{3}{4}$, if the latter is a non-negative integer. Hence this quantity must be equal to $(n-1)$, where $n = 1, 2, 3, \cdots$

$$\frac{\lambda}{4} - \frac{l}{2} - \frac{3}{4} = n - 1, \quad \text{with } n = 1, 2, 3, \cdots . \tag{12.25}$$

Since $v(\zeta) = L^{l+\frac{1}{2}}_{\frac{\lambda}{4}+\frac{l}{2}-\frac{1}{4}}(\zeta)$ is a polynomial in ζ of degree $(n-1)$, it is a polynomial in r of degree $2(n-1)$. Using Eq. (12.24) complete radial part of the eigen function is

$$\frac{\tilde{R}(r)}{r} = C_{nl}\ (\alpha r)^l\ L^{l+\frac{1}{2}}_{\frac{\lambda}{4}+\frac{l}{2}-\frac{1}{4}}(\alpha^2 r^2)\ \mathrm{e}^{-\frac{1}{2}\alpha^2 r^2}. \tag{12.26}$$

The normalization constant C_{nl} and the series expansion of the associated Laguerre polynomial can be found in Chap. 7 of Ref. Roy and Nigam (1967).

The product of first two r-dependent terms of the wave function, Eq. (12.26) is a polynomial in r of degree $l + 2(n-1) = N$ (say), where N is an integer ≥ 0, since l is an integer ≥ 0. Substituting for λ from Eq. (12.21) in Eq. (12.25) we get the energy eigen value

$$\begin{aligned} E_{nl} \equiv E &= \hbar\omega_c\Big[2(n-1) + l + \frac{3}{2}\Big] = \hbar\omega_c\Big(N + \frac{3}{2}\Big), \quad N = 0, 1, 2, \cdots , \\ &\text{where } N = 2n + l - 2, \\ &\text{with } n = 1, 2, 3, \cdots \text{ and } l = 0, 1, 2, \cdots . \end{aligned} \tag{12.27}$$

Thus the energy eigen value is the same as in the treatment of the problem in Cartesian coordinates [see Eq. (12.18)], as it should.

Degeneracy: Since energy eigen value is independent of the magnetic quantum number m for a given l, there is a degeneracy factor $(2l+1)$ for each l. This is due to the rotational symmetry about the z-axis for all spherically symmetric potentials in general. In addition E_{nl} is the same for different combinations of (n, l), for which $2n + l - 2 = N$ has the same value. Now for a given N, the quantity $2n = N + 2 - l$ must be an even integer ≥ 2. Hence for even (or odd) N, only even (or odd) l values contribute and the degenerate eigen functions correspond to the (n, l) combinations [Roy and Nigam (1967)]

$$(n,l) = \Big(\frac{N+2}{2},0\Big), \Big(\frac{N}{2},2\Big), \cdots, \Big(2,N-2\Big), \Big(1,N\Big) \text{ for } N = \text{ even},$$
$$(n,l) = \Big(\frac{N+1}{2},1\Big), \Big(\frac{N-1}{2},3\Big), \cdots, \Big(2,N-2\Big), \Big(1,N\Big) \text{ for } N = \text{ odd}.$$

Then the degeneracy for a given N is

$$D(N) = \sum_{k=0}^{\frac{N}{2}} \Big[2(2k)+1\Big] \text{ for } N,l \text{ even, with } l = 2k,\ k = 0,1,2,\cdots,$$
$$D(N) = \sum_{k=0}^{\frac{N-1}{2}} \Big[2(2k+1)+1\Big] \text{ for } N,l \text{ odd, with } l = 2k+1,\ k = 0,1,2,\cdots$$

The result is the same for both cases, and we have

$$D(N) = \frac{1}{2}(N+1)(N+2), \quad \text{for } N = 0,1,2,\cdots. \tag{12.28}$$

This agrees with Eq. (12.20) for the same value of N (as it should), although the two treatments are entirely different.

12.5 Problems

1. Consider a quantum mechanical particle with energy less than the height V_0 of the barrier, in a cylindrical hole with *permeable* walls. Take the potential to be zero inside the cylinder and V_0 (finite) outside it. Obtain a transcendental equation for energy of the particle in terms of parameters of the potential.
2. Calculate the expectation values of r and r^2 for the ground and first excited states of a spherically symmetric oscillator, using Eq. (12.26). Use Ref. Roy and Nigam (1967) for normalization constant and series expansion of associated Laguerre polynomials.

References

Abramowitz, M., Stegun, I.A. (eds.): Handbook of Mathematical Functions. Dover Publications Inc., New York (1972)
Arfken, G.: Mathematical Methods for Physicists. Academic Press, New York (1966)
Chattopadhyay, P.K.: Mathematical Physics. New Age International (P) Ltd., New Delhi (2006)
Roy, R.R., Nigam, B.P.: Nuclear Physics: Theory and Experiment. John Wiley & Sons, New York (1967)

Chapter 13
Scattering in One Dimension



Abstract In this chapter one-dimensional scattering situations, namely a free particle encountering a rigid wall, penetration through a finite square barrier and scattering of a free particle by a delta function barrier, have been discussed.

Keywords Free particle motion · Rigid wall · Penetration through finite square barrier · Delta barrier

We next consider unbound states of quantum systems. In such states, a particle can be found even at an infinite distance. Hence the wave function does not vanish at infinity, but must be finite. Such a situation can occur if a particle comes from a great distance, interacts with a time-independent localized potential (i.e. the potential vanishes at infinity) and moves again to a great distance. This process is called *scattering* of the particle by the potential. As we discussed in Chap. 8, the particle can have any positive energy, since it can be found at infinity. Indeed this is a time-dependent problem, but the motion is governed by the time-independent Schrödinger equation for time-independent potentials. For simplicity, we first consider one-dimensional problems in this chapter. We will consider a more realistic three-dimensional situation in the next chapter.

Since the wave function extends to infinity, it needs special normalization, as we discussed in Chap. 3, Sect. 5. In some cases only the relative amplitude of various parts of the wave functions is needed, as we see below. We consider some simple cases in this chapter.

13.1 A Free Particle Encountering an Infinitely Rigid Wall

Consider a particle of mass m moving in $+x$-direction in the region $x < 0$ and encountering an infinitely rigid wall at $x = 0$. The potential for $x < 0$ can be taken as zero. Hence

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5_13

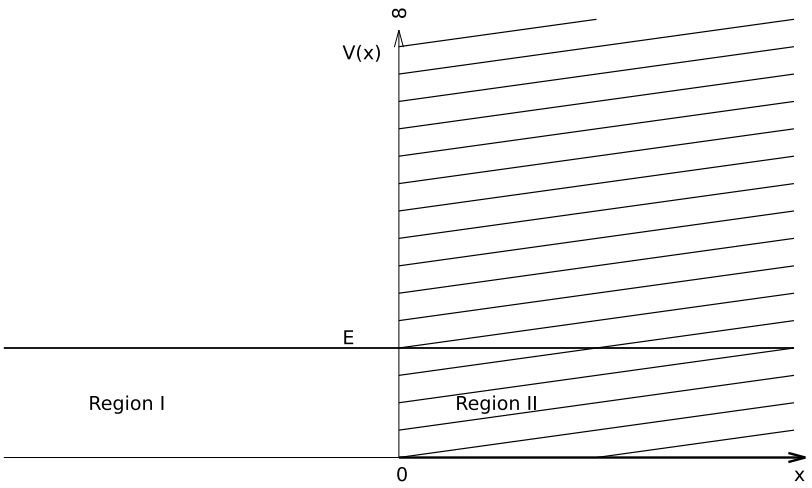


Fig. 13.1 Infinite potential barrier at $x = 0$ in one dimension. Cross-hatched region indicates $V = +\infty$

$$\begin{aligned} V(x) &= 0, \quad -\infty < x < 0 \quad \text{(Region I)} \\ &= \infty, \quad x \geq 0. \qquad\quad \text{(Region II)} \end{aligned}$$

Figure 13.1 shows the potential. The Schrödinger equationin Region I is

$$-\frac{\hbar^2}{2m}\frac{d^2u_{\text{I}}(x)}{\mathrm{d}x^2} = Eu_{\text{I}}(x) \;\Rightarrow\; \frac{d^2u_{\text{I}}(x)}{\mathrm{d}x^2} + k^2u_{\text{I}}(x) = 0,$$
$$\text{where } k = +\sqrt{\frac{2mE}{\hbar^2}}. \quad \text{(Region I)}.$$

In Region II, $V(x) = \infty$, hence $u_{\text{II}}(x) = 0$ since for any finite value of $u_{\text{II}}(x)$ in this region, Vu_{II} becomes infinite and the Schrödinger equation cannot be satisfied. Physically this means that the particle with a finite E cannot penetrate the rigid wall. The general solution in Region I is

$$u_{\text{I}}(x) = A\mathrm{e}^{ikx} + B\mathrm{e}^{-ikx},$$

where A and B are constants. As both the functions are finite as $x \to -\infty$, both of them survive. The complete time-dependent wave function is

$$\psi_{\text{I}}(x,t) = u_{\text{I}}(x)\mathrm{e}^{-i\omega t} = A\mathrm{e}^{i(kx-\omega t)} + B\mathrm{e}^{-i(kx+\omega t)}, \quad (\text{where } \omega = \frac{E}{\hbar}). \qquad (13.1)$$

The first and second terms represent plane waves traveling in $+x$- and $-x$-directions respectively. The wave function at the infinite potential discontinuity must vanish (see

Chap. 8, Sect. 8.1). Hence the boundary condition at $x = 0$ is $\psi_{\text{I}}(x = 0, t) = 0$. Then from Eq. (13.1)

$$(A + B)\mathrm{e}^{-i\omega t} = 0 \;\Rightarrow\; B = -A.$$

Hence

$$\psi_{\text{I}}(x, t) = A\left[\mathrm{e}^{i(kx-\omega t)} - \mathrm{e}^{-i(kx+\omega t)}\right]. \tag{13.2}$$

The final expression is a superposition of two plane waves traveling in $+x$- and $-x$-directions *with equal amplitudes*. This results in a standing wave [$\propto \sin(kx)\mathrm{e}^{-i\omega t}$]. The first term on the right side of Eq. (13.2) represents the wave *incident* on the infinite barrier from the left, and the second is the *reflected* wave. The reflection coefficient is defined as

$$\mathcal{R} = \left|\frac{\text{amplitude of reflected wave}}{\text{amplitude of incident wave}}\right|^2 = \left|\frac{-A}{A}\right|^2 = 1,$$

independent of the normalization constant. This shows that a particle incident on the rigid barrier is *reflected totally* by the barrier, with no transmission. Momentum is $p = \hbar k$ and $-p = -\hbar k$ for the particle moving to the right and to the left respectively. Thus it is discontinuous at $x = 0$, the difference in momentum is absorbed by the infinite barrier. This agrees with the classical result.

13.2 Penetration Through a Finite Square Barrier

We next consider penetration through a finite barrier – a square barrier for simplicity [following Ref. Schiff (1955)]

$$\begin{aligned} V(x) &= 0 \quad \text{for } -\infty < x < 0 \text{ (Region I) and } x > a \text{ (Region III)},\\ &= V_0 \quad \text{for } 0 \le x \le a \text{ (Region II)}, \end{aligned}$$

with $V_0 > 0$.

Figure 13.2 shows the potential. Suppose a particle of mass m and energy E traveling in Region I to the right from $x \to -\infty$ is incident on the barrier at $x = 0$. According to quantum mechanics, we expect the particle with any energy $E > 0$ to be partially reflected back to Region I and partially transmitted into Region III. Due to this lack of symmetry, we do not gain anything by choosing the barrier symmetrically about the origin. The Schrödinger equation in Regions I and III is

$$-\frac{\hbar^2}{2m}\frac{d^2u(x)}{\mathrm{d}x^2} = Eu(x) \;\Rightarrow\; \frac{d^2u(x)}{\mathrm{d}x^2} + k^2u(x) = 0 \quad \text{where } k = +\sqrt{\frac{2mE}{\hbar^2}},$$

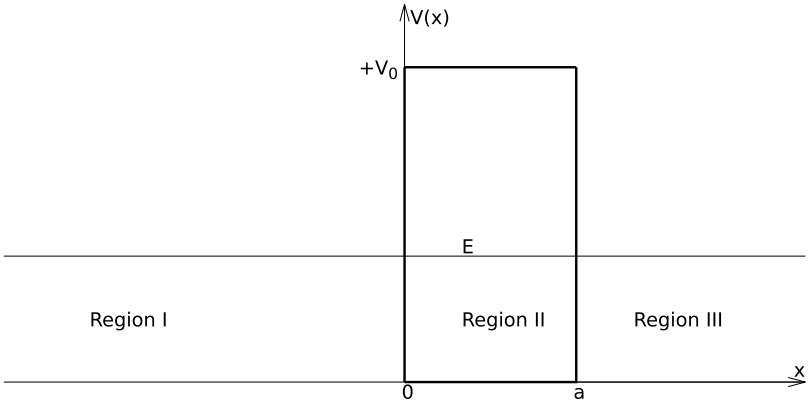


Fig. 13.2 Finite potential barrier of height V_0 and width a in one dimension

whose solutions are

$$\begin{aligned} u_{\text{I}}(x) &= A\mathrm{e}^{ikx} + B\mathrm{e}^{-ikx} \quad \text{in region I,} \\ u_{\text{III}}(x) &= C\mathrm{e}^{ikx} \qquad\qquad\quad \text{in Region III.} \end{aligned} \tag{13.3}$$

We do not have any wave going to the left in Region III. As before, the complete wave function, including the time-dependent part $\mathrm{e}^{-i\omega t}$ (with $\omega = \frac{E}{\hbar}$) is

$$\begin{aligned} u_{\text{I}}(x) &= A\mathrm{e}^{i(kx-\omega t)} + B\mathrm{e}^{-i(kx+\omega t)} \quad \text{in region I,} \\ u_{\text{III}}(x) &= C\mathrm{e}^{i(kx-\omega t)} \qquad\qquad\quad \text{in Region III.} \end{aligned}$$

For Region I, we retain both waves traveling to the right and to the left (corresponding respectively to the incident and reflected waves), while only the wave traveling to the right in Region III (corresponding to the transmitted wave). The reflection ($\mathcal{R}$) and transmission ($\mathcal{T}$) coefficients are given by

$$\mathcal{R} = \left|\frac{B}{A}\right|^2, \quad \text{and} \quad \mathcal{T} = \left|\frac{C}{A}\right|^2.$$

The normalization constants can be obtained by any one of the three procedures mentioned in Chap. 3, Sect. 5. However, since we need only the ratios, we will leave them as they are. The Schrödinger equation in Region II is

$$-\frac{\hbar^2}{2m}\frac{d^2u(x)}{\mathrm{d}x^2} + V_0 u(x) = Eu(x),$$

which becomes

$$\frac{d^2u_{\text{II}}(x)}{\mathrm{d}x^2} + K^2 u_{\text{II}}(x) = 0 \quad \text{where} \quad K = +\sqrt{\frac{2m(E-V_0)}{\hbar^2}} \quad \text{for} \quad E > V_0,$$
$$\frac{d^2u_{\text{II}}(x)}{\mathrm{d}x^2} - \beta^2 u_{\text{II}}(x) = 0 \quad \text{where} \quad \beta = +\sqrt{\frac{2m(V_0-E)}{\hbar^2}} \quad \text{for} \quad E < V_0.$$

Solution of these equations are

$$u_{\text{II}}(x) = F\mathrm{e}^{iKx} + G\mathrm{e}^{-iKx} \qquad \text{for} \quad E > V_0,$$
$$u_{\text{II}}(x) = F'\mathrm{e}^{\beta x} + G'\mathrm{e}^{-\beta x} \qquad \text{for} \quad E < V_0. \tag{13.4}$$

For the time being, we leave the time dependence and apply the conditions of continuity of $u(x)$ and $\frac{\mathrm{d}u(x)}{\mathrm{d}x}$ at $x = 0$ and $x = a$. These lead to *four* equations involving *five* unknown constants (*viz.* A, B, F, G, C for $E > V_0$ and A, B, F', G', C for $E < V_0$). However if we divide the entire wave function given by Eqs. (13.3) and (13.4) by A (we are permitted to do this, as this corresponds to a new overall normalization constant), then we have four equations for four unknowns $\frac{B}{A}$, $\frac{F}{A}$, $\frac{G}{A}$ and $\frac{C}{A}$ for $E > V_0$ and similarly for $E < V_0$. Define a dimensionless quantity $\epsilon = \frac{E}{V_0}$

For $E > V_0$ $(\epsilon > 1)$

From the continuity conditions at $x = 0$ and $x = a$, we can eliminate $\frac{F}{A}$ and $\frac{G}{A}$ to get $\frac{B}{A}$ and $\frac{C}{A}$ for $E > V_0$, and calculating the absolute squares of these, we have after some algebraic steps [see Ref. Schiff (1955)]

$$\mathcal{R} = \left|\frac{B}{A}\right|^2 = \left[1 + \frac{4k^2K^2}{(k^2-K^2)^2\sin^2(Ka)}\right]^{-1}$$
$$\mathcal{T} = \left|\frac{C}{A}\right|^2 = \left[1 + \frac{(k^2-K^2)^2\sin^2(Ka)}{4k^2K^2}\right]^{-1}$$

Expressing K and k in terms of E and V_0 and then the dimensionless quantity ϵ, we have

$$\mathcal{R} = \left[1 + \frac{4\epsilon(\epsilon-1)}{\sin^2(Ka)}\right]^{-1} \quad \text{for} \quad \epsilon > 1, \tag{13.5}$$

and

$$\mathcal{T} = \left[1 + \frac{\sin^2(Ka)}{4\epsilon(\epsilon-1)}\right]^{-1} \quad \text{for} \quad \epsilon > 1. \tag{13.6}$$

We can easily verify that

$$\mathcal{R} + \mathcal{T} = 1 \quad \Rightarrow \quad |B|^2 + |C|^2 = |A|^2,$$

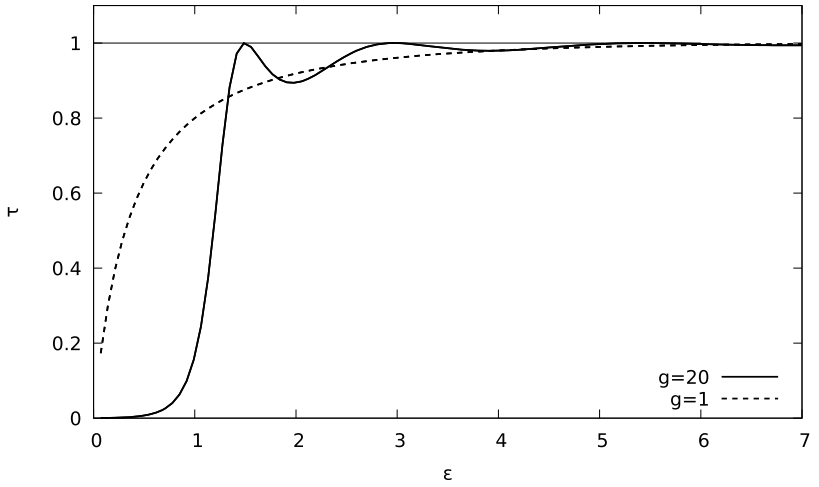


Fig. 13.3 Transmission coefficient ($\mathcal{T}$) as a function of ϵ for two different values of g, *viz.*, $g = 1$ (fairly transparent barrier) and $g = 20$ (quite opaque barrier)

which corresponds to the conservation of probability. The sum of reflection and transmission coefficients is unity, meaning that there is no loss or gain of particle fluxes.

For $0 < E < V_0 \;\; (0 < \epsilon < 1)$

In this case K is replaced by

$$K \xrightarrow[E<V_0]{} \sqrt{-\frac{2m(V_0 - E)}{\hbar^2}} = i\beta,$$

and

$$\sin(Ka) = \sin(i\beta a) = i \sinh(\beta a).$$

Using these in Eqs. (13.5) and (13.6), we have

$$\mathcal{R} = \left[1 + \frac{4\epsilon(1-\epsilon)}{\sinh^2(\beta a)}\right]^{-1} \quad \text{for} \;\; 0 < \epsilon < 1, \tag{13.7}$$

and

$$\mathcal{T} = \left[1 + \frac{\sinh^2(\beta a)}{4\epsilon(1-\epsilon)}\right]^{-1} \quad \text{for} \;\; 0 < \epsilon < 1. \tag{13.8}$$

We introduce another dimensionless quantity g as

$$g = \frac{2mV_0a^2}{\hbar^2},$$

and express the dimensionless quantities Ka and βa as

$$Ka = \sqrt{g(\epsilon - 1)} \quad \text{and} \quad \beta a = \sqrt{g(1 - \epsilon)}.$$

In Fig. 13.3 we plot $\mathcal{T}$ against ϵ for $g = 1$ and $g = 20$, using Eqs. (13.8) for $\epsilon < 1$ and (13.6) for $\epsilon > 1$. Taking the limit $\epsilon \to 1$ from both sides, we can verify that $\mathcal{T}$ is continuous at $\epsilon = 1$. For $\epsilon > 1$, we see that $\sin(Ka) = 0$ for $Ka = n\pi$ with $n = 1, 2, \cdots$. Hence $\mathcal{T} = 1$ and there is complete transmission. Thus whenever the barrier contains an integral number of half wave lengths ($\frac{\lambda}{2} = \frac{\pi}{K}$), there is complete transmission (this occurs when the outside wave function matches smoothly with the wave function within the barrier, for $\epsilon > 1$). Thus the nth complete transmission occurs at $\epsilon = 1 + \frac{n^2\pi^2}{g}$. The first one (for $n = 1$) is at $\epsilon = 1.4935$ for $g = 20$, while at $\epsilon = 10.8696$ for $g = 1$. A look at Fig. 13.3 shows that $\mathcal{T}$ increases monotonically up to the point where it reaches the first maximum ($= 1$), then oscillates with decreasing amplitude, touching $\mathcal{T} = 1$ at $\epsilon = 1 + \frac{n^2\pi^2}{g}$. For $\epsilon \gg 1$, Eq. (13.6) shows that $\mathcal{T}$ approaches 1. This phenomenon is analogous to the interference of light passing through a refracting plate.

13.3 Scattering of a Free Particle by a δ-Barrier

Suppose a particle of mass m and energy $E > 0$ is incident from left on a δ-function barrier at $x = a$

$$V(x) = C\delta(x - a),$$

with C and a positive constants. Note that C has dimension of energy$\times$length, as δ-function is defined through $\int \delta(x - a)\mathrm{d}x = 1$. We name the regions $x < a$ and $x > a$ as Regions I and II respectively. The Schrödinger equation is

$$-\frac{\hbar^2}{2m}\frac{d^2u}{\mathrm{d}x^2} + C\delta(x - a)u = Eu.$$

For $x \neq a$, we have

$$\frac{d^2u}{\mathrm{d}x^2} + \alpha^2 u = 0, \quad \text{where } \alpha = \sqrt{\frac{2mE}{\hbar^2}}.$$

In Region I, we have the incident wave going to the right (amplitude A) and a reflected wave going to the left (amplitude B'), hence

$$u_{\mathrm{I}}(x) = A\mathrm{e}^{i\alpha x} + B'\mathrm{e}^{-i\alpha x}.$$

In Region II, there is only the transmitted wave going to the right (amplitude F'), hence

$$u_{\text{II}}(x) = F' \mathrm{e}^{i\alpha x}.$$

The first derivative of the wave function is discontinuous at the infinite potential step at $x = a$, while the wave function itself is continuous (see Chap. 8). Continuity of wave function at $x = a$ gives

$$A\mathrm{e}^{i\alpha a} + B'\mathrm{e}^{-i\alpha a} = F'\mathrm{e}^{i\alpha a}$$

Dividing by A and substituting $B = \frac{B'}{A}$ and $F = \frac{F'}{A}$, we have

$$\mathrm{e}^{i\alpha a} + B\mathrm{e}^{-i\alpha a} = F\mathrm{e}^{i\alpha a}. \tag{13.9}$$

Discontinuity of $u'(x)$ across the δ-function is obtained by integrating the Schrödinger equation from $a - \epsilon$ to $a + \epsilon$ and then taking the limit $\epsilon \to 0$. This gives

$$\left.\frac{\mathrm{d}u_{\text{II}}}{\mathrm{d}x}\right|_a - \left.\frac{\mathrm{d}u_{I}}{\mathrm{d}x}\right|_a - C_1 u_{\text{II}}(a) = 0, \qquad \text{where } C_1 = \frac{2mC}{\hbar^2}.$$

Note that both C_1 and α have dimension $[L^{-1}]$. Substituting for u_I and u_{II} and dividing by $i\alpha A$, we get

$$\mathrm{e}^{i\alpha a} - B\mathrm{e}^{-i\alpha a} = (1 + \frac{iC_1}{\alpha})F\mathrm{e}^{i\alpha a}. \tag{13.10}$$

Solving Eqs. (13.9) and (13.10) for B and F, we have

$$B = \left[\frac{-iC_1}{2\alpha + iC_1}\right]\mathrm{e}^{2i\alpha a}, \quad F = \left[\frac{2\alpha}{2\alpha + iC_1}\right].$$

Reflection ($\mathcal{R}$) and transmission ($\mathcal{T}$) coefficients are

$$\mathcal{R} = \left|\frac{B'}{A}\right|^2 = |B|^2 = \frac{C_1^2}{4\alpha^2 + C_1^2}, \quad \mathcal{T} = \left|\frac{F'}{A}\right|^2 = |F|^2 = \frac{4\alpha^2}{4\alpha^2 + C_1^2}.$$

Expressing α^2 in terms of E and C_1 in terms of C, we have

$$\mathcal{R} = \frac{C^2}{C^2 + \frac{2E\hbar^2}{m}}, \quad \mathcal{T} = \frac{\frac{2E\hbar^2}{m}}{C^2 + \frac{2E\hbar^2}{m}}. \tag{13.11}$$

Note that both $\mathcal{R}$ and $\mathcal{T}$ are dimensionless, as they should. From Eq. (13.11), we see that $R + T = 1$, as expected. Also for $E = 0$, $R = 1$, $T = 0$, while for $C = 0$ (i.e. no barrier) $T = 1$, $R = 0$, as our intuition suggests.

13.4 Problems

1. Consider a one-dimensional barrier of finite height V_0 (> 0) but of infinite extent from $x = 0$ to $x = \infty$. A free particle of energy $E > 0$ is incident from the left. Calculate the reflection coefficient for both $E < V_0$ and $E > V_0$ and comment on the results.
2. A free particle of energy E, with $0 < E < V_0$, is incident from the left on the one-dimensional potential barrier

$$\begin{aligned} V(x) &= \frac{1}{2}bx^2, \quad -a \leq x \leq a, \\ &= 0 \qquad\quad \text{otherwise}, \end{aligned}$$

where $V_0 = \frac{1}{2}ba^2$. Using the standard solutions of the Hermite equation, write down the continuity conditions. Obtain an expression for the transmission coefficient.

Reference

Schiff, L.I.: Quantum Mechanics, 2nd edn. McGraw-Hill Book Company Inc. (1955); reprinted as International Student Edition, Kogakusha Co. Ltd. Tokyo

Chapter 14
Scattering in Three Dimension

Abstract Three-dimensional scattering has been discussed, clearly mentioning the separate (and apparently conflicting) idealizations in laboratory setup and theoretical analysis. These are justified by specifying the widely different scales of length, mass, etc. This discussion provides understanding of both the experimental setup and the theoretical analysis, making a convincing bridge between the two. Partial waves, phase shift, etc., for spherically symmetric potentials are presented. Coulomb scattering, Green's function in scattering, Born approximation, resonance scattering are also included.

Keywords Idealizations in theoretical treatment · Kinematics · Scattering cross section · Spherically symmetric potential · Partial waves · Optical theoren · Phase shift · Sign of phase shift · Ramsauer-Townsend effect · Rigid sphere scattering · Coulomb scattering · Green's function · Born approximation · Resonance scattering

In this chapter we will discuss the quantum mechanical problem of scattering in three dimensions. As in classical mechanics, an incoming particle is scattered by a fixed force field or two particles collide with each other, interacting due to their mutual force. For simplicity we will initially consider only forces of finite range. This force may be produced by another particle at the origin, the force being the result of interaction between the particles. This will be a two-body collision. An electron scattered by another electron is an example of this type. Two-body collisions are important to have direct information about their mutual interaction. In a potential scattering, the force on the incident particle is produced by a *fixed* potential field, centered at the origin, leading to a one-body scattering by a force field. An example of this type is an electron scattered by an external electromagnetic field.

As in Chap. 11, Sect. 11.1, one can separate the two-body collision with *mutual interaction* into a relative motion and the free motion of the center of mass (CM). The former is the motion of a single fictitious particle of reduced mass μ moving in a potential field $V(\vec{r})$, where $\vec{r}$ is the relative separation. The CM motion is of no interest for studying the system of two particles. This is similar to the bound state problem of two mutually interacting particles (as in the H-atom).

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5_14

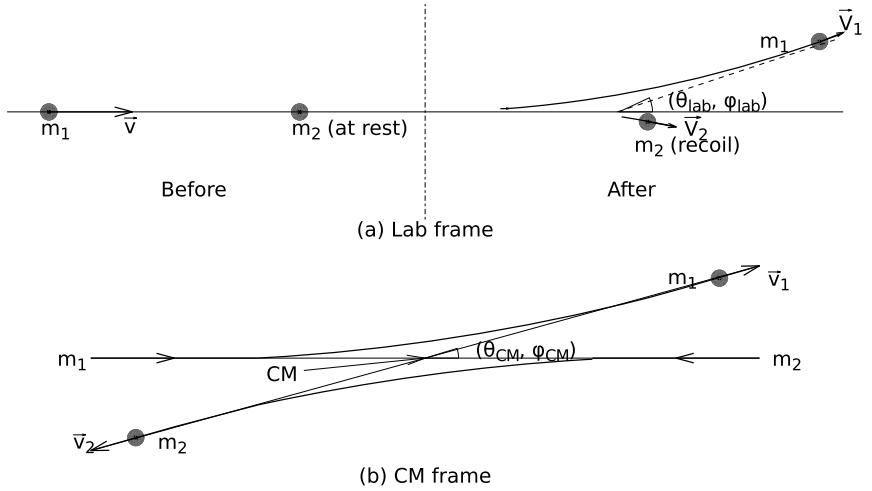


Fig. 14.1 Scattering of a particle of mass m_1 by another of mass m_2 in lab frame (top) and CM frame (bottom)

Experimental arrangement: In the actual experimental situation, the target particle is usually at rest and the bombarding particle (called projectile) is moving. During the scattering ("collision"), the incident particle feels the force due to the target particle and moves in general in a direction other than the incident direction and the target particle recoils. Recoil of the target particle within the target material is small and not usually detected. In this process energy and momentum are conserved. The coordinate frame in which the target particle is initially at rest is called the *laboratory system* or lab frame (lab) and the coordinate frame in which the center of mass is at rest – initially and always (due to conservation of linear momentum, as there is no net force on the center of mass) – is called the *center of mass* (CM) system or CM frame (see Fig. 14.1).

For the theoretical analysis, we have to keep in mind the experimental arrangement (see Fig. 14.2). A single individual particle is usually not detected in an experimental situation (because of the experimental difficulty of producing a single particle and the extremely small probability of its detection). An *incident* (usually parallel) *beam* of particles is made to impinge on a "*target material*" containing a number of target particles, so that the net current of scattered particles in the detector is large enough to be accurately measured. Thus it is indeed a many-body process. However, flux of the incident beam (as well as the scattered beam) is small enough, so that the average instantaneous separation between any two particles is very large compared with the range of interaction (if any) between the incident particles. Similarly the number of target particles in the "target" is large enough for the detector to detect the net current (of all particles scattered by many target particles) with reasonable precision, while it is small enough so that the average separation of target particles is much larger than the range of force. Furthermore, thickness of the target material should be small to avoid multiple scattering within this material. Thin target material

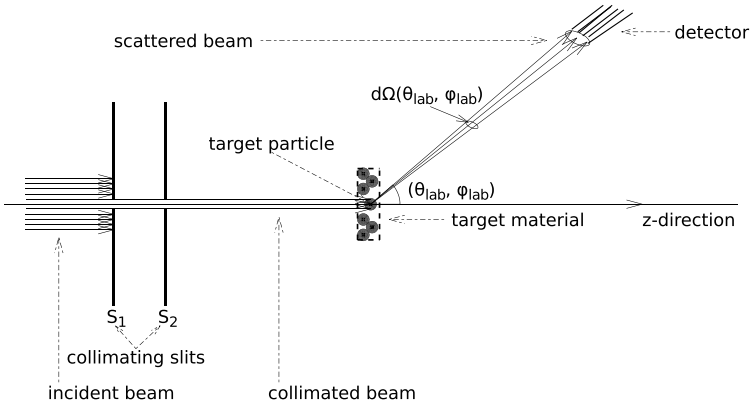


Fig. 14.2 Schematic experimental setup for scattering in lab frame. Note that the dimensions in the figure and number of particles in the target are only indicative and do not represent the actual situation

and a narrow incident beam also help in fairly well defined scattering angles. The scattered particles are detected by a detector holding a *small* solid angle $d\Omega_{lab}$ (see below for meaning of "smallness" of $d\Omega_{lab}$) about the direction $(\theta_{lab}, \phi_{lab})$ at the origin. Hence for the theoretical analysis, we can consider a beam of single particles being scattered by *a single target particle* (see Fig. 14.3). For further simplification, this can be considered as the scattering of a single particle moving in $+z$-direction by another single particle at rest at the origin (see Fig. 14.1). This simplification is possible for the vastly different length scales of macroscopic and quantum systems. The net current of scattered particles as recorded by the detector (and also the current of incident particles, measured separately) gives the quantum mechanical probability of scattering (see below).

Another important point to note is the following. The incident beam is collimated (by slits or by electromagnetic fields for charged particles) to a small cross-sectional area [say, a few $(\text{mm})^2$]. This is necessary, so that the detector placed a few tens of cm from the target in a direction $(\theta_{lab}, \phi_{lab})$ does not record the incident particles directly (see Fig. 14.2). The cross-sectional area of the detector should be small enough so that the solid angle $d\Omega_{lab}$ subtended by it is small enough for a fairly accurate $(\theta_{lab}, \phi_{lab})$ measurement, but large enough for the detector to record a large number of scattered particles. Thus the experimental arrangement is a tricky one, which has to be cleverly designed, balancing various aspects. It is clear that for the scattering cross-section (i.e., the probability of scattering) we have to measure incident flux and scattered flux, instead of individual particles. The laboratory scattering can be envisaged as *a large ensemble of independent two-body scattering by a time-independent potential*, at different times. This agrees with probabilistic interpretation of quantum mechanics.

The laboratory arrangements described above are for standard low energy scattering experiments. Special arrangements can be made for experiments with very few particles, as in high energy physics experiments. Also note that the scattering process is indeed time-dependent, involving scattering of traveling wave packets representing

the incident and the target particles. For a time-independent potential, we simplify (see below) by solving the time-independent Schrödinger equation for the relative motion of a mutually interacting two-body system. Furthermore, for simplicity we will consider elastic scattering by spherically symmetric potentials only, although some general statements will be made where necessary.

14.1 Kinematics for Scattering

We will consider non-relativistic motion and solve the Schrödinger equation for the relative motion of the interacting particles. For kinematics, we will use classical dynamics. This is justified, since in the experimental setup the particles are well localized and classical mechanics is satisfied for such well localized wave packets (see Chap. 8, Sect. 8.6). We follow Ref. Schiff (1968) for the following treatment, with simplifications mentioned in the introduction.

Clearly velocities, scattering angles and scattering cross-sections (probability of scattering) as measured in the laboratory (the "lab-system") and in the center-of-mass frame (the "CM-system") will differ. These are experimentally measured in the lab-system, while the theoretical analysis will be done for the relative motion in the CM-system. In the lab-system, a particle is incident on the target particle which is at rest. Hence the CM moves also in the same direction. To find relations between velocities, angles, etc., in the two systems, we can translate the lab-system, in the direction of incidence, so that the CM is at rest. In the lab-system, consider a particle of mass m_2 at rest at the origin and another particle of mass m_1 incident along $+\hat{z}$-direction on m_2 with speed v (see Fig. 14.1, top). After scattering, m_1 moves in the direction $(\theta_{\text{lab}}, \phi_{\text{lab}})$ (called the scattering angles in the lab-system) with speed V_1, while m_2 recoils with speed V_2 in another direction. Only the particle m_1 is detected in the detector placed a few tens of cm from the origin.

In the CM-system (Fig. 14.1, bottom) the particles move in opposite directions, both before and after the collision. In this system, both m_1 and m_2 approach the CM (which is fixed and taken as the origin) in opposite directions along z-axis. After the collision m_1 moves in the direction $(\theta_{\text{CM}}, \phi_{\text{CM}})$, which are the scattering angles in the CM-system. Hence m_2 moves after the collision in the directly opposite direction $(\pi - \theta_{\text{CM}}, \pi + \phi_{\text{CM}})$.

We consider spherically symmetric potentials $(V(\vec{r}) = V(r))$ which do not affect azimuthal (ϕ) motion. Hence the scattering takes place in a plane, so that

$$\phi_{\text{lab}} = \phi_{\text{CM}}.$$

Next, let speed of the CM in lab frame be v' and speeds of m_1 and m_2 in the CM-system be v_1 and v_2 (in opposite directions). Then the relative speed $= v_1 + v_2 =$ speed of m_1 in lab $= v$. Also CM is defined to have net momentum zero (therefore it should more appropriately be called center of momentum system). Hence the magnitudes of momentum of m_1 and m_2 must be the same in the CM-system, both before and after the collision (they move in opposite directions)

$$m_1 v_1 = m_2 v_2,$$
$$v_1 + v_2 = v.$$

Solving we get

$$v_1 = \frac{m_2 v}{m_1 + m_2}, \qquad v_2 = \frac{m_1 v}{m_1 + m_2}. \tag{14.1}$$

We will be considering *elastic scattering*, for which the initial and final total kinetic energy of the two particles, before and after scattering, are the same. In *inelastic scattering* these are different due to the conversion of internal energy into kinetic energy and *vice versa*. Since in the CM-system, the momenta of the two particles must always be equal and opposite, we see that after the scattering m_1 and m_2 recede with speeds v_1 and v_2 respectively for elastic scattering. Resolving the velocity of the outgoing m_1 along $\hat{z}$-direction and along its perpendicular in the plane of scattering, we have

$$\begin{aligned} v_1 \cos\theta_{\text{CM}} + v' &= V_1 \cos\theta_{\text{lab}} \\ v_1 \sin\theta_{\text{CM}} &= V_1 \sin\theta_{\text{lab}}. \end{aligned}$$

In the above, $\vec{V}_1$ and $\vec{V}_2$ are the velocities of m_1 and m_2 in the lab-system *after the collision*. Dividing the second relation by the first one above, we have

$$\tan\theta_{\text{lab}} = \frac{v_1 \sin\theta_{\text{CM}}}{v_1 \cos\theta_{\text{CM}} + v'} = \frac{\sin\theta_{\text{CM}}}{\cos\theta_{\text{CM}} + (\frac{v'}{v_1})}. \tag{14.2}$$

Now speed of CM (v') in lab frame is obtained by the requirement that the momentum *of the CM* (due to its motion in lab frame) is the same as the total momentum of the particles in lab

$$(m_1 + m_2)v' = m_1 v + 0.m_2 \qquad \text{(for elastic scattering)}.$$

Therefore

$$v' = \frac{m_1 v}{(m_1 + m_2)}. \tag{14.3}$$

From Eqs. (14.1) and (14.3).

$$\frac{v'}{v_1} = \frac{m_1}{m_2} \equiv \gamma \qquad \text{(for elastic scattering)}. \tag{14.4}$$

Here γ is the ratio of the speed of CM in the lab frame to the speed of m_1 in CM frame. The value of γ will be different for inelastic scattering [see Eq. (14.11)]. Then from Eq. (14.2),

$$\begin{aligned} \tan\theta_{\text{lab}} &= \frac{\sin\theta_{\text{CM}}}{\cos\theta_{\text{CM}} + (\frac{m_1}{m_2})} = \frac{\sin\theta_{\text{CM}}}{\cos\theta_{\text{CM}} + \gamma} \\ \phi_{\text{lab}} &= \phi_{\text{CM}}. \end{aligned} \tag{14.5}$$

Now

$$\begin{aligned} E_{\text{lab}} &= \frac{1}{2} m_1 v^2, \\ E_{\text{CM}} &= \frac{1}{2} m_1 v_1^2 + \frac{1}{2} m_2 v_2^2 \\ &= \frac{1}{2} \frac{m_1 m_2}{m_1 + m_2} v^2 = \frac{1}{2} \mu v^2, \end{aligned} \tag{14.6}$$

where $\mu = \frac{m_1 m_2}{m_1 + m_2}$ is the reduced mass. E_{CM} is the kinetic energy of the two particles in the CM frame. In an elastic scattering it remains the same before and after the scattering. Thus

$$E_{\text{CM}} = \frac{m_2}{m_1 + m_2} E_{\text{lab}} \qquad \text{(for elastic scattering).} \tag{14.7}$$

Thus if $m_2 >> m_1$, E_{CM} is almost the same as E_{lab}. But if $m_1 >> m_2$, CM energy is much less than the lab energy. Even if $m_1 = m_2$, only half of lab energy is available for the CM scattering. In high energy experiments, the technical difficulty and the cost increase very rapidly as the energy of the incident particle is increased. Hence experiments (involving for example proton-proton scattering) are designed in *colliding beam accelerators*, in which both the particles are accelerated in opposite directions, so that the lab frame coincides with the CM frame.

14.2 Scattering Cross-Section

In the laboratory, scattering of a single incident particle by a single target particle is not possible, except in special devices (like cloud chamber and photographic plates). This is so, for the difficulty of producing and identifying a single particle, as also the difficulty of detection of a single particle. Moreover, to get the scattering cross-section (probability of scattering) we should have a large number of such systems (ensembles). Usually, a beam of incident particles is scattered by a target material, containing a number of target particles. The target material may be a thin solid film (or a thin container holding a liquid or a gas) containing the target particle. Provisions are to be made for unwanted objects in the target material.

Consider a parallel beam of incident particles having N particles per unit area per unit time (i.e., flux $= N$) bombarding a target containing n particles (see Fig. 14.2). Let the number of particles scattered within a small solid angle $\mathrm{d}\Omega_{\text{lab}}$, centered about the direction $(\theta_{\text{lab}}, \phi_{\text{lab}})$ per unit time be N_{sc}. Then clearly N_{sc} is proportional to N, n and $\mathrm{d}\Omega_{\text{lab}}$ or

$$N_{\text{sc}} = Nn \left(\frac{\mathrm{d}\sigma(\theta_{\text{lab}}, \phi_{\text{lab}})}{\mathrm{d}\Omega} \right)_{\text{lab}} \mathrm{d}\Omega_{\text{lab}}. \tag{14.8}$$

where the proportionality constant $\left(\frac{d\sigma(\theta_{lab},\phi_{lab})}{d\Omega}\right)_{lab}$ is called the differential scattering cross-section in the lab frame. We see from Eq. (14.8), that the differential scattering cross-section is the number of particles scattered per unit time by a single scatterer per unit incident flux per unit solid angle. In Eq. (14.8), N_{sc}, N, n, and $d\Omega_{lab}$ respectively have dimensions T^{-1}, $(L^{-2}T^{-1})$, dimensionless, and steradian (unit of solid angle). Hence $\frac{d\sigma}{d\Omega}$ has dimension area/steradion. The total lab scattering cross-section, defined as

$$\sigma_{lab} = \int \left(\frac{d\sigma}{d\Omega}(\theta_{lab}, \phi_{lab})\right)_{lab} d\Omega_{lab} \equiv \int \left(\frac{d\sigma}{d\Omega}\right)_{lab} \sin\theta_{lab}\, d\theta_{lab}\, d\phi_{lab}, \tag{14.9}$$

will have dimension of area (standard unit for nuclear systems is $10^{-24}\text{cm}^2 = 1$ barn). This is why σ is called 'cross-section'. In Eq. (14.9) and in the following, we suppress dependence of $\frac{d\sigma}{d\Omega}$ on polar angles for brevity, indicating only the system (lab- or CM-system) as its subscript.

The differential scattering cross-section in the CM-system is obtained by the condition that (by definition) the number of particles scattered in lab through $d\Omega_{lab}$ about $(\theta_{lab}, \phi_{lab})$ is the same as the number of particles in CM scattered through corresponding $d\Omega_{CM}$ about (θ_{CM}, ϕ_{CM}). Then from Eq. (14.8)

$$\left(\frac{d\sigma}{d\Omega}\right)_{lab} \sin\theta_{lab}\, d\theta_{lab}\, d\phi_{lab} = \left(\frac{d\sigma}{d\Omega}\right)_{CM} \sin\theta_{CM}\, d\theta_{CM}\, d\phi_{CM}.$$

Using Eq. (14.5) we get [Schiff (1968)]

$$\left(\frac{d\sigma}{d\Omega}\right)_{lab} = \frac{(1+\gamma^2+2\gamma\cos\theta_{CM})^{\frac{3}{2}}}{|1+\gamma\cos\theta_{CM}|}\left(\frac{d\sigma}{d\Omega}\right)_{CM}, \tag{14.10}$$

where $\gamma = \frac{m_1}{m_2}$ for elastic scattering. Equations (14.5) and (14.10) are valid also for inelastic scattering with γ replaced by [Schiff (1968)]

$$\gamma = +\left(\frac{m_1 m_3}{m_2 m_4}\,\frac{E_{CM}}{E_{CM}+Q}\right)^{\frac{1}{2}}, \tag{14.11}$$

where E_{CM} is the initial kinetic energy (in the CM-system) of the two particles *before* scattering. After an inelastic scattering, two new particles with masses m_3 and m_4 emerge and an amount of energy Q is converted from internal potential energy (due to interaction between the particles) to the kinetic energy of the emergent particles. The process is called *exothermic* or *endothermic* according to $Q > 0$ or $Q < 0$ respectively. Equation (14.10) is valid for elastic scattering also, for which $m_3 = m_1$, $m_4 = m_2$ and $Q = 0$.

14.3 Schrödinger Equation

We will do the theoretical analysis in the CM-system. However the experiment is done in the lab-system. Using relations (14.5), (14.7) and (14.10) quantities in the lab-system can be obtained in terms of respective calculated quantities in the CM-system. These are then compared with the experimentally measured quantities in the lab-system. Since our theoretical analysis will be in the CM-system only, henceforth we drop the suffix CM in all quantities.

Here we list some of the simplifications in the theoretical treatment *vis-à- vis* the experimental procedure. The experimental scattering process is a time-dependent, many-body process. But since individual particles in both the incident and scattered beams are well separated in quantum length scale (their wave packets do no overlap), we can solve the Schrödinger equation for only one incident particle and one target particle. Since the interaction between them is assumed mutual, the relevant Schrödinger equation becomes the one for relative motion. We also assume the potential to be time-independent, and hence we can solve the time-independent Schrödinger equation for the mutual interaction. Finally we assume (for simplicity) the potential to be spherically symmetric, for which the scattering takes place in a single plane and there is no ϕ dependence. Hence we will drop the ϕ variable later on, after some general relations.

In Chap. 11, Sect. 1, we saw that the Schrödinger equation of two mutually interacting particles can be separated into: (a) Schrödinger equation for the relative motion in CM frame and (b) Schrödinger equation for the CM, moving as a free particle. We are interested here in (a) only, satisfying

$$-\frac{\hbar^2}{2\mu}\nabla_{\vec{r}}^2\,\psi(\vec{r}\,)+V(\vec{r}\,)\,\psi(\vec{r}\,)=E\psi(\vec{r}\,), \tag{14.12}$$

where

$$\vec{r}=\vec{r_1}-\vec{r_2}\ \text{(relative separation)},\quad \text{and}\quad \mu=\frac{m_1m_2}{m_1+m_2}\ \text{(reduced mass)}.$$

$E\ (=E_{\rm CM})$ is the CM energy. From Eq. (14.6) we have,

$$E=\frac{1}{2}\mu v^2.$$

Thus we can consider Eq. (14.12) as describing the elastic collision of a fictitious particle of mass μ, initial speed $v=\frac{\hbar k}{\mu}$ and kinetic energy $E=\frac{1}{2}\mu v^2$, by a potential field $V(\vec{r}\,)$ located at a fixed scattering center at the origin. Its interaction is described by the potential $V(\vec{r}\,)$, where $\vec{r}$ is the vector separation of the fictitious particle from the origin.

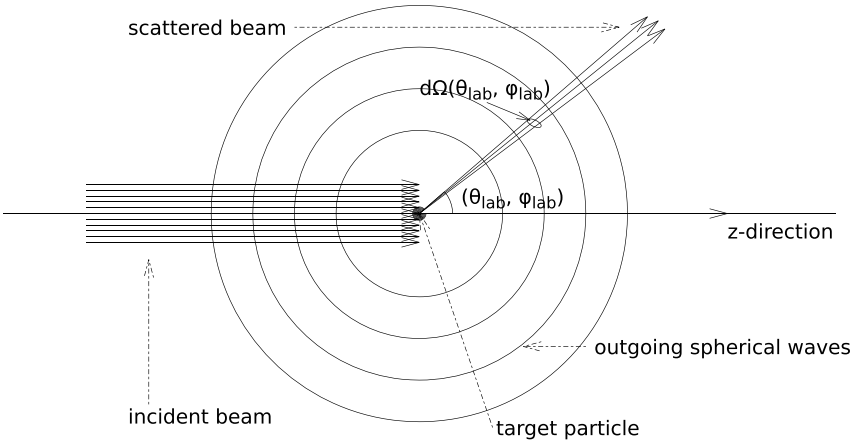


Fig. 14.3 Idealized scattering by a single target particle in the CM frame

We will first consider the potential $V(\vec{r})$ to be of finite range, so that the incident particle, when still far off from the scattering center, will be unaffected by $V(\vec{r})$ and it can be taken as an infinite plane wave (since it is moving in a particular direction, which we take as $+\hat{z}$-direction). See Fig. 14.3. Then the wave function of the incident wave is a plane wave propagating in $\hat{z}$ direction

$$\psi_{\text{in}}(\vec{r}) = A\mathrm{e}^{ikz} = A\mathrm{e}^{ikr\cos\theta}.$$

Normalization:
Since $\psi_{in}(\vec{r})$ does not represent a bound state, $\int |\psi_{in}(\vec{r})|^2 d^3r$ diverges. Hence we cannot normalize by equating the normalization integral to unity. We can however use one of several possibilities, according to Chap. 3, Sect. 5:

1. We can normalize to unit incident flux. If we calculate the flux due to $\psi_{\text{in}}(\vec{r})$, we get

$$\begin{aligned}\vec{j}_{\text{in}}(\vec{r}) &= \frac{\hbar}{2i\mu}\left[\psi^*_{in}(\vec{r})\left(\vec{\nabla}_{\vec{r}}\,\psi_{\text{in}}(\vec{r})\right) - \left(\vec{\nabla}_{\vec{r}}\,\psi^*_{\text{in}}(\vec{r})\right)\psi_{\text{in}}(\vec{r})\right] \\ &= |A|^2\,\frac{\hbar k}{\mu}\,\hat{z} = |A|^2 v\,\hat{z}. \end{aligned} \tag{14.13}$$

   Then normalization to unit flux gives $A = \frac{1}{\sqrt{v}}$.
2. Box normalization: We can imagine the particle confined within a large cubical box of sides L aligned parallel to the Cartesian axes, subject to periodic boundary conditions. Thus we require $\int_{\text{cube of sides L}} |\psi_{\text{in}}(\vec{r})|^2 d^3r = 1$ together with periodic boundary conditions. This gives $A = \frac{1}{L^{\frac{3}{2}}}$.
3. We can use δ-function normalization, for which

$$A = \frac{1}{(2\pi)^{\frac{3}{2}}}.$$

We will choose A after adding the scattered part in the following.

Once the incident wave comes within range of interaction between it and the target, there will be generation of scattered wave $\psi_{\text{sc}}(\vec{r})$, which are spherical waves outgoing from the CM (see Fig. 14.3), i.e., the origin and has the from $\propto \frac{e^{ikr}}{r}$, with the wave number $k = \sqrt{\frac{2\mu E}{\hbar^2}}$. Now due to the interaction, the amplitude of the scattered wave $\psi_{\text{sc}}(\vec{r})$ will be angle-dependent (i.e., different in different directions), hence

$$\psi_{\text{sc}}(\vec{r}) \propto f(\theta, \phi)\frac{e^{ikr}}{r},$$

where the amplitude $f(\theta, \phi)$ of ψ_{sc} is called the 'scattering amplitude'. The asymptotic form of the complete solution of Eq. (14.12) will be the superposition of $\psi_{\text{in}}(\vec{r})$ and $\psi_{\text{sc}}(\vec{r})$

$$\psi(\vec{r}) \underset{r\to\infty}{\longrightarrow} \psi_{\text{in}}(\vec{r}) + \psi_{\text{sc}}(\vec{r}) = A\Big[e^{ikz} + f(\theta, \phi)\frac{e^{ikr}}{r}\Big].$$

We take the normalization constants of $\psi(\vec{r})$ and $\psi_{\text{in}}(\vec{r})$ the same, since in actual case, the probability of scattering is very small (see later), $|\psi_{\text{sc}}(\vec{r})| << |\psi_{\text{in}}(\vec{r})|$ and the two normalization constants are practically the same. Moreover, the interference term between these parts can be disregarded for calculation of incidence and scattered currents.

We can visualize the situation by a simple analogy: consider a tranquil pond with a placid surface, in which there is a small floating object at the center. Now suppose a continuous train of plane traveling surface waves is generated at a distance. As the plane wave strikes the floating object, the latter starts vibrating and generating circular outgoing surface waves (compare with Fig. 14.3). These two interfere and produce a pattern on the surface. The quantum mechanical process is a three-dimensional analogue of this scenario, except that the scattered waves in the macroscopic analogue have an appreciable amplitude as the scatterer (floating object) is also macroscopic. The amplitude of quantum mechanical scattered wave from a *single scatterer* is much smaller than that of the incident wave. Another important difference is that in the experimental setup for scattering, the incident wave is screened from the detector by collimating the incident wave to a narrow beam.

Now $\psi_{\text{in}}(\vec{r})$ satisfies the asymptoticSchrödinger equation , i.e., Equation (14.12) with $V(\vec{r}) = 0$. We can verify that $\psi_{\text{sc}}(\vec{r})$ also satisfies this equation, up to terms of order $O(\frac{1}{r})$. Since (we drop suffix $\vec{r}$ of $\nabla^2_{\vec{r}}$, where no confusion arises)

$$\begin{aligned}\nabla^2 &= \frac{1}{r^2}\frac{\partial}{\partial r}(r^2\frac{\partial}{\partial r}) + \frac{1}{r^2}\Big[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\Big(\sin\theta\frac{\partial}{\partial\theta}\Big) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\Big]\\ &= \frac{1}{r^2}\frac{\partial}{\partial r}\Big(r^2\frac{\partial}{\partial r}\Big) - \frac{\hat{L}^2}{r^2\hbar^2},\end{aligned}$$

we have up to terms of $O(\frac{1}{r})$.

$$\nabla^2\Big[f(\theta,\phi)\frac{e^{ikr}}{r}\Big] = f(\theta,\phi)\frac{1}{r^2}\frac{\partial}{\partial r}\Big(r^2\frac{\partial}{\partial r}(\frac{e^{ikr}}{r})\Big) - \frac{1}{r^3}\frac{e^{ikr}}{\hbar^2}\hat{L}^2 f(\theta,\phi).$$

The first term is

$$\frac{f(\theta,\phi)}{r^2}\frac{\partial}{\partial r}\Big[r^2\Big(\frac{ike^{ikr}}{r} - \frac{e^{ikr}}{r^2}\Big)\Big] = -k^2 f(\theta,\phi)\frac{e^{ikr}}{r}.$$

The second term goes as $\frac{1}{r^3}$. Keeping terms of $O(\frac{1}{r})$, we have

$$-\frac{\hbar^2}{2\mu}\nabla^2\psi_{sc} = \frac{\hbar^2k^2}{2\mu}f(\theta,\phi)\frac{e^{ikr}}{r} = E\psi_{sc}.$$

Hence ψ_{sc} is an asymptotic (hence $V(\vec{r}\,) \to 0$) solution of Eq. (14.12), up to leading powers of $\frac{1}{r}$. Thus $\psi = \psi_{in} + \psi_{sc}$ is also a solution of Eq. (14.12) in the asymptotic region, $V = 0$. Since final cross-section is independent of the normalization, we will simply take $A = 1$

$$\psi(\vec{r}) \underset{r\to\infty}{\longrightarrow} e^{ikz} + f(\theta,\phi)\frac{e^{ikr}}{r}, \quad (k = \sqrt{\frac{2\mu E}{\hbar^2}}). \tag{14.14}$$

As mentioned earlier, we disregard the interference term between $\psi_{in}(\vec{r}\,)$ and $\psi_{sc}(\vec{r}\,)$ for calculation of the incident and scattered currents, as also for the overall normalization. Its justification follows.

In the experimental setup, the incident beam is well collimated by slits S_1 and S_2 (say) with a cross-sectional area of a few $(\text{mm})^2$. This is experimentally important, so that the error in θ_{lab} measurement is minimized (see Fig. 14.2). Thus the incident beam (beyond the slits S_1, S_2) is not a pure infinite plane wave, but a truncated one. Its transverse extent ($\sim$ a few mm) is so large compared to quantum length scale ($\sim 10^{-8}$ cm for atomic and $\sim 10^{-13}$ cm for nuclear systems) that the wave can be treated as practically plane. On the other hand the detector is placed at a distance of a few tens of cm from the scatterer. Thus, for not too small scattering angles, the incident beam cannot reach the detector and the current seen by the detector is due to the scattered wave only. The scattered current density is

$$\vec{j}_{sc} = \frac{\hbar}{2i\mu}\Big[\psi^*_{sc}(\vec{r}\,)\Big(\vec{\nabla}_{\vec{r}}\,\psi_{sc}(\vec{r}\,)\Big) - \Big(\vec{\nabla}_{\vec{r}}\,\psi^*_{sc}(\vec{r}\,)\Big)\psi_{sc}(\vec{r}\,)\Big]. \tag{14.15}$$

Now

$$\vec{\nabla}_{\vec{r}} = \hat{r}\frac{\partial}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial}{\partial\theta} + \hat{\phi}\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi} \equiv \hat{r}\frac{\partial}{\partial r} + \frac{1}{r}\vec{\nabla}_{\theta,\phi}.$$

Then

$$\begin{aligned}\vec{j}_{\rm sc} &= \frac{\hbar}{2i\mu}\Big[\Big\{f^*\frac{\mathrm{e}^{-ikr}}{r}\Big(f\frac{\partial}{\partial r}\Big(\frac{\mathrm{e}^{ikr}}{r}\Big)\Big) - \Big(f^*\frac{\partial}{\partial r}\Big(\frac{\mathrm{e}^{-ikr}}{r}\Big)\Big).f\frac{\mathrm{e}^{ikr}}{r}\Big\}\hat{r} \\ &\quad + \frac{1}{r^3}\Big\{f^*\Big(\vec{\nabla}_{\theta,\phi}f\Big) - \Big(\vec{\nabla}_{\theta,\phi}f^*\Big)f\Big\}\Big] \\ &= \frac{\hbar k}{\mu}\frac{|f|^2}{r^2}\hat{r} + O\left(\frac{1}{r^3}\right) \underset{r\to\infty}{\longrightarrow} \frac{v|f(\theta,\phi)|^2}{r^2}\hat{r} \quad \left(\text{up to } O\left(\frac{1}{r^2}\right)\right).\end{aligned}$$

Hence at a sufficiently large distance from the scatterer [actually a few tens of cm in experimental setup, which is very large compared to atomic ($\sim 10^{-8}$ cm) or nuclear ($\sim 10^{-13}$ cm) length scales], $\vec{j}_{\rm sc}$ is dominantly radial and has a magnitude $\frac{v|f(\theta,\phi)|^2}{r^2}$. Then the number of particles scattered through $\mathrm{d}\Omega$ about the direction (θ, ϕ) is

$$N_{\rm sc} = \vec{j}_{\rm sc}\,.\,\mathrm{d}\vec{A} \underset{r\to\infty}{\longrightarrow} \frac{v|f(\theta,\phi)|^2}{r^2}\,\hat{r}\,\cdot\, r^2\mathrm{d}\Omega\,\hat{r} = v|f(\theta,\phi)|^2\mathrm{d}\Omega.$$

We have from the definition of scattering cross-section in CM frame [equivalent to Eq. (14.8) for CM frame, with the suffix 'CM' dropped in our present notation], with incident flux $N = v$ (see Eq. (14.13) with our chosen $A = 1$), and a single scatterer $(n = 1)$

$$N_{\rm sc} = Nn\Big(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\Big)\mathrm{d}\Omega = v.1.\Big(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\Big)\mathrm{d}\Omega = v|f(\theta,\phi)|^2\mathrm{d}\Omega, \qquad \text{from above.}$$

Equating we get the differential scattering cross-section

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f(\theta,\phi)|^2. \tag{14.16}$$

Note that for spherically symmetric potentials, there is no ϕ dependence. An integration over all solid angles gives the total scattering cross-section σ

$$\sigma = \int |f(\theta,\phi)|^2\mathrm{d}\Omega.$$

14.4 Spherically Symmetric Potential: Method of Partial Waves

We consider $V(\vec{r}\,) = V(r)$, i.e., $V(\vec{r}\,)$ depends on the scalar r, but not on (θ, ϕ). Such a potential is called spherically symmetric or central (since the force is always along $\hat{r}$). Moreover we assume that $V(\vec{r}\,)$ is of finite range, i.e., for $r \to \infty$, $V(\vec{r}\,) = 0$. This is an important class of potentials found in nature. The Schrödinger equation is

$$\Big[-\frac{\hbar^2}{2\mu}\nabla_{\vec{r}}^2 + V(r) - E\Big]\psi(\vec{r}\,) = 0. \tag{14.17}$$

For a complete spherically symmetric problem, a general solution would contain spherical harmonics in the wave function (see Chap. 10). However in the scattering problem the spherical symmetry is broken due to the incident beam in the $\hat{z}$-direction. But the azimuthal symmetry about this axis is preserved. Consequently, although l is not preserved, m is. Since there is no dependence on ϕ , we can take $m = 0$. Hence we can expand $\psi(\vec{r}\,)$ in the complete set of Legendre polynomials, $\{P_l(\cos\theta)\}$ (see Chap. 7, Sect. 7.1)

$$\psi(\vec{r}\,) = \sum_{l=0}^{\infty} R_l(r)\,P_l(\cos\theta). \tag{14.18}$$

$R_l(r)$ is called the lth partial wave. When there is no azimuthal symmetry, we need to expand in the compete set of spherical harmonics $\{Y_{lm}(\theta,\phi)\}$ instead of $\{P_l(\cos\theta)\}$. Substitution of Eq. (14.18) in Eq. (14.17) gives [compare with Eq. (10.11) with eigen value $\lambda = l(l+1)$ of $\hat{L}^2$ operator]

$$\frac{1}{r^2}\frac{d}{dr}\Big(r^2\frac{dR_l}{dr}\Big) + \Big[k^2 - U(r) - \frac{l(l+1)}{r^2}\Big]R_l(r) = 0,$$
$$\text{with } k = \sqrt{\frac{2\mu E}{\hbar^2}},\quad U(r) = \frac{2\mu V(r)}{\hbar^2}. \tag{14.19}$$

Note that as $r \to \infty$, $V(r)$ and hence $U(r) \to 0$. Boundary condition at $r = 0$ requires $R_l(r)$ to be finite (so that $\psi(\vec{r}\,)$ is finite). The boundary condition at $r \to \infty$ is that $\psi(\vec{r}\,)$ must have the form (14.14) [note that there is no ϕ dependence]

$$\psi(\vec{r}\,) \underset{r\to\infty}{\longrightarrow} e^{ikz} + f(\theta)\frac{e^{ikr}}{r},$$

which will be imposed to determine the normalization of $R_l(r)$ and to calculate $f(\theta)$. For a finite range potential the asymptotic form of Eq. (14.19) becomes

$$\frac{d^2R_l}{dr^2} + \frac{2}{r}\frac{dR_l}{dr} + [k^2 - \frac{l(l+1)}{r^2}]R_l(r) = 0.$$

This is just the spherical Bessel equation, Eq. (7.69), if we multiply by r^2 and change variable from r to $x = kr$

$$x^2 j_l''(x) + 2x j_l'(x) + [x^2 - l(l+1)]j_l(x) = 0.$$

Hence a general solution is a linear combination of spherical Bessel and spherical Neumann functions. Note that we are studying the asymptotic solution and hence

there is no restriction to include $n_l(kr)$ which diverges only at $r = 0$. Thus

$$R_l(r) \underset{r\to\infty}{\longrightarrow} A_l\, j_l(kr) + B_l\, n_l(kr) \equiv C_l\Big[\cos\delta_l\, j_l(kr) - \sin\delta_l\, n_l(kr)\Big], \tag{14.20}$$

where A_l, B_l are constants. New constants C_l and δ_l are defined through

$$A_l = C_l\cos\delta_l, \quad B_l = -C_l\sin\delta_l.$$

Next note the asymptotic expansions for large r [see Ref. Arfken (1966)]

$$j_l(kr) \underset{r\to\infty}{\longrightarrow} \frac{1}{kr}\sin(kr - l\frac{\pi}{2}),$$
$$n_l(kr) \underset{r\to\infty}{\longrightarrow} -\frac{1}{kr}\cos(kr - l\frac{\pi}{2}). \tag{14.21}$$

Then from Eqs. (14.20) and (14.21)

$$R_l(r) \underset{r\to\infty}{\longrightarrow} \frac{C_l}{kr}\Big[\sin(kr - l\frac{\pi}{2})\cos\delta_l + \cos(kr - l\frac{\pi}{2})\sin d_l\Big] = \frac{C_l}{kr}\sin(kr - l\frac{\pi}{2} + \delta_l). \tag{14.22}$$

Thus from Eqs. (14.22) and (14.18)

$$\psi(\vec{r}) \underset{r\to\infty}{\longrightarrow} \sum_{l=0}^{\infty}\frac{C_l}{kr}\sin(kr - l\frac{\pi}{2} + \delta_l)P_l(\cos\theta) = \sum_{l=0}^{\infty}\frac{C_l}{2ikr}\Big[e^{i(kr - l\frac{\pi}{2} + \delta_l)} - e^{-i(kr - l\frac{\pi}{2} + \delta_l)}\Big]P_l(\cos\theta), \tag{14.23}$$

δ_l is called *phase shift* for the lth partial wave, which will be discussed later. We next compare this with the asymptotic form (14.14) of the total wave function identified earlier. The partial wave expansion of the plane wave $(\psi_{\text{in}}(\vec{r}))$ has the form [see Ref. Arfken (1966)]

$$e^{ikz} = e^{ikr\cos\theta} = \sum_{l=0}^{\infty}(2l+1)i^l\, j_l(kr)\, P_l(\cos\theta) \underset{r\to\infty}{\longrightarrow} \sum_{l=0}^{\infty}\frac{(2l+1)i^l}{kr}\sin(kr - l\frac{\pi}{2})P_l(\cos\theta) = \sum_{l=0}^{\infty}\frac{(2l+1)i^l}{2ikr}\Big[e^{i(kr - l\frac{\pi}{2})} - e^{-i(kr - l\frac{\pi}{2})}\Big]P_l(\cos\theta). \tag{14.24}$$

Substituting Eq. (14.24) in (14.14) and equating with Eq. (14.23) we have

$$\sum_{l=0}^{\infty} \frac{(2l+1)i^l}{2ikr}\left[e^{i(kr-l\frac{\pi}{2})} - e^{-i(kr-l\frac{\pi}{2})}\right]P_l(\cos\theta) + f(\theta)\frac{e^{ikr}}{r}$$
$$= \sum_{l=0}^{\infty} \frac{C_l}{2ikr}\left[e^{i(kr-l\frac{\pi}{2}+\delta_l)} - e^{-i(kr-l\frac{\pi}{2}+\delta_l)}\right]P_l(\cos\theta).$$

Since e^{+ikr} and e^{-ikr} are linearly independent, we equate the coefficients of e^{-ikr} and e^{+ikr} on both sides to get

$$-\sum_{l=0}^{\infty} \frac{(2l+1)i^l}{2ikr} e^{il\frac{\pi}{2}} P_l(\cos\theta) = -\sum_{l=0}^{\infty} \frac{C_l}{2ikr} e^{il\frac{\pi}{2}} e^{-i\delta_l} P_l(\cos\theta),$$
$$\sum_{l=0}^{\infty} \frac{(2l+1)i^l}{2ikr} e^{-il\frac{\pi}{2}} P_l(\cos\theta) + f(\theta)\frac{1}{r} = \sum_{l=0}^{\infty} \frac{C_l}{2ikr} e^{-il\frac{\pi}{2}} . e^{i\delta_l} P_l(\cos\theta).$$

Since these are true for all values of θ and $\{P_l(\cos\theta)\}$ forms an orthogonal set, we have from the first of above relations

$$C_l = (2l+1)i^l e^{i\delta_l},$$

and from the second, after substituting for C_l, we have

$$f(\theta) = \frac{1}{2ik}\sum_{l=0}^{\infty}[C_l e^{i\delta_l} - (2l+1)i^l]e^{-il\frac{\pi}{2}} P_l(\cos\theta)$$
$$= \frac{1}{2ik}\sum_{l=0}^{\infty}(2l+1)i^l(e^{2i\delta_l} - 1)e^{-il\frac{\pi}{2}} P_l(\cos\theta)$$
$$= \frac{1}{2ik}\sum_{l=0}^{\infty}(2l+1)(e^{2i\delta_l} - 1)P_l(\cos\theta) \quad \text{(because } i^l e^{-il\frac{\pi}{2}} = i^l(-i)^l = +1\text{)}$$
$$= \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)e^{i\delta_l}\sin\delta_l P_l(\cos\theta). \tag{14.25}$$

Then the differential scattering cross-section is

$$\frac{d\sigma}{d\Omega} = |f(\theta)|^2 = \frac{1}{k^2}\left|\sum_{l=0}^{\infty}(2l+1)e^{i\delta_l}\sin\delta_l P_l(\cos\theta)\right|^2. \tag{14.26}$$

Total elastic scattering cross-section, (σ) is obtained by integrating $\frac{d\sigma}{d\Omega}$ over all angles

$$\sigma = \int \frac{d\sigma}{d\Omega} d\Omega = \frac{1}{k^2} \int_0^{\pi} \int_0^{2\pi} \sum_{l,l'=0}^{\infty} (2l+1)(2l'+1) e^{i\delta_l} e^{-i\delta_{l'}} \sin\delta_l \sin\delta_{l'}$$
$$\times P_l(\cos\theta) P_{l'}(\cos\theta) \sin\theta d\theta d\phi$$
$$= \frac{2\pi}{k^2} \sum_{l,l'=0}^{\infty} (2l+1)(2l'+1) e^{i(\delta_l - \delta_{l'})} \sin\delta_l \sin\delta_{l'} \int_{-1}^{1} P_l(x) P_{l'}(x) dx$$
$$= \frac{4\pi}{k^2} \sum_{l=0}^{\infty} (2l+1) \sin^2\delta_l \equiv \sum_{l=0}^{\infty} \sigma_l, \tag{14.27}$$

using the orthonormality of Legendre polynomials

$$\int_{-1}^{1} P_l(x) P_{l'}(x) dx = \frac{2}{2l+1} \delta_{ll'}.$$

In Eq. (14.27), σ_l is called total cross-section of the lth partial wave.

14.4.1 Optical Theorem

From Eq. (14.25)

$$f(0) = \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) e^{i\delta_l} \sin\delta_l P_l(1)$$
$$= \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) e^{i\delta_l} \sin\delta_l \qquad (\text{since } P_l(1) = 1)$$
$$= \frac{1}{k} \sum_{l=0}^{\infty} (2l+1)(\cos\delta_l + i \sin\delta_l) \sin\delta_l.$$

Hence

$$I_m f(0) = \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) \sin^2\delta_l = \frac{k}{4\pi}\sigma \quad [\text{from Eq. (14.27)}].$$
$$\text{Thus } \sigma = \frac{4\pi}{k} I_m f(0). \tag{14.28}$$

This is called the optical theorem. It is called "optical theorem" because of its analogy with absorption of light passing through a medium. The imaginary part of the complex refractive index is related to the absorption coefficient, i.e., the total absorption cross-section. We can understand Eq. (14.28) in the following fashion. The process of scattering removes particles from the incident beam by an amount proportional to σ (total scattering cross-section). Thus beam current is smaller in the forward direction ($\theta \simeq 0$) than in the backward direction ($\theta \simeq \pi$). This is produced by interference between the two terms in the asymptotic expression

$$\psi(\vec{r}\,) \underset{r\to\infty}{\longrightarrow} e^{ikz} + f(\theta)\frac{e^{ikr}}{r}.$$

The interference term is a linear function of the forward scattering amplitude $[f(0)]$. This justifies a relation of the general form of Eq. (14.28). Actually this relation holds much more generally: when f depends on ϕ also, as well as when σ includes inelastic and elastic scattering [see Refs. Schiff (1968), Merzbacher (1965)].

14.4.2 Phase Shifts

δ_l introduced in the first of Eq. (14.23), *viz*

$$\psi(\vec{r}\,) \underset{r\to\infty}{\longrightarrow} \sum_{l=0}^{\infty} \frac{C_l}{kr} \sin\left(kr - l\frac{\pi}{2} + \delta_l\right) P_l(\cos\theta),$$

is the phase difference of the lth partial wave between the asymptotic form of actual radial wave function $R_l(r)$ [Eq. (14.22] and the asymptotic form of the wave function in the absence of scattering potential, *viz*, Eq. (14.24)

$$e^{ikz} \underset{r\to\infty}{\longrightarrow} \sum_{l=0}^{\infty} \frac{(2l+1)i^l}{kr} \sin\left(kr - l\frac{\pi}{2}\right) P_l(\cos\theta).$$

We saw that δ_l completely determines the scattering cross-section [see Eqs. (14.26) and (14.27)]. σ vanishes when $\delta_l = 0°$ or $180°$ for all l. Our derivation of scattering amplitude $f(\theta)$ [Eq. (14.25)] is valid even if a radius a does not exist, beyond which $V(r)$ is negligible. It is valid provided $V(r)$ falls off more rapidly than $1/r$, since we kept terms of order $O(\frac{1}{r})$ in our treatment. However, the partial wave method is particularly useful if the potential is of a finite range a. In that case, for a given energy of the incident particle, only a few l-values ($\leq l_{\max}$) contribute to Eqs. (14.25), (14.26) or (14.27). We can make an estimate of $l_{\max}$ as follows.

Suppose the impact parameter (which is the perpendicular distance of m_2 from the line of incidence of m_1, see Fig. 14.4) be b. Then semiclassically

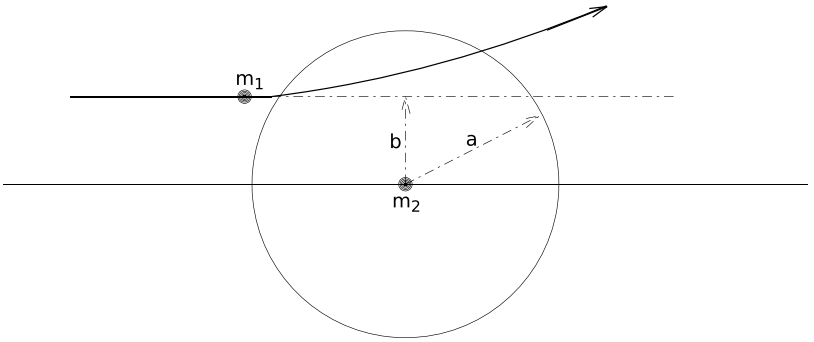


Fig. 14.4 Impact parameter (b) and range of interaction (a) for scattering of m_1 by m_2. Scattering takes place if $b < a$

$$|\vec{L}| = p.b = \hbar kb \simeq \hbar\sqrt{l(l+1)},$$
$$\text{Therefore} \qquad b = \frac{\sqrt{l(l+1)}}{k}.$$

Here we take the semiclassical value of L to be $\hbar\sqrt{l(l+1)}$, since all components of $\vec{L}$ cannot be simultaneously measured, but $\hat{L}^2$ has a value $\hbar^2 l(l+1)$.

If the interaction has a finite range a, then the motion of m_1 will be unaffected if $b > a$, i.e., if

$$\frac{\sqrt{l(l+1)}}{k} > a. \tag{14.29}$$

Hence the lth partial wave will be unaffected (i.e., $\delta_l = 0$) if

$$E_{\text{lab}} = \frac{\hbar^2 k^2}{2m_1} < \frac{\hbar^2 l(l+1)}{2m_1 a^2}.$$

Putting $l = 1$, we see that only $l = 0$ partial wave will be scattered if

$$E_{\text{lab}} < \frac{\hbar^2 l(l+1)}{2m_1 a^2} = \frac{\hbar^2}{m_1 a^2}.$$

For example, for neutron-proton (n-p) scattering $\frac{\hbar^2}{m_1} = 41.47$ Mev fm^2 and $a \simeq 2$ fm, as the range of nuclear force is about 2 fm (note that 1 fm $= 10^{-13}$ cm). Therefore for $E_{\text{lab}} \leq 10$ Mev, only $l = 0$ partial waves will be affected in the n-p scattering.

Calculation of δ_l

The phase shift is computed by equating log derivative of the asymptotic radial solution, Eq. (14.20), with that of the radial wave function $[(R_l)_{\text{in}}]$ obtained by actually

solving the complete Schrödinger equation in the inside region ($r < a$) at $r = a$. Suppose

$$\frac{1}{(R_l)_{in}} \frac{d(R_l)_{in}}{\mathrm{d}r}\Big|_{r=a} = \gamma_l. \tag{14.30}$$

where $(R_l)_{\text{in}}$ is the inside solution for $r < a$. Then

$$\begin{aligned} \gamma_l &= \frac{1}{(R_l)} \frac{d(R_l)}{\mathrm{d}r}\Big|_{r=a} \qquad \text{where } R_l \text{ is given by Eq. (14.20)} \\ &= \frac{k[\cos\delta_l j_l'(ka) - \sin\delta_l n_l'(ka)]}{[\cos\delta_l j_l(ka) - \sin\delta_l n_l(ka)]}. \end{aligned}$$

This gives

$$\tan\delta_l = \frac{kj_l'(ka) - \gamma_l j_l(ka)}{kn_l'(ka) - \gamma_l n_l(ka)}. \tag{14.31}$$

Even if the potential does not vanish for $r > a$ we can use Eq. (14.31) for a sufficiently large $r = a$, provided $V(r)$ decreases faster than $\frac{1}{r}$. Note that δ_l depends on energy, since both k and γ_l depend on energy.

14.4.3 *Relation Between Sign of Phase Shift (δ_l) and the Nature (Attractive or Repulsive) of Potential*

To find a relation between the sign of phase shift and the nature of potential we proceed as follows [Sengupta (2003)]. The radial Schrödinger equation (14.19) is

$$\frac{1}{r^2}\frac{d}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R_l(r)}{\mathrm{d}r}\right) + \left[k^2 - U(r) - \frac{l(l+1)}{r^2}\right]R_l(r) = 0.$$

Putting $R_l(r) = \frac{\phi_l(r)}{r}$, it becomes

$$\frac{d^2\phi_l(r)}{\mathrm{d}r^2} + \left[k^2 - U(r) - \frac{l(l+1)}{r^2}\right]\phi_l(r) = 0. \tag{14.32}$$

The boundary condition at $r = 0$ is $R_l(0) =$ finite. Hence $\phi_l(0) = 0$. Now, replace $U(r)$ by $\lambda U(r)$, in which $U(r)$ is positive (purely repulsive) and λ is a real constant, which can be positive or negative. Then the potential is repulsive or attractive accordingly as λ is positive or negative. We have

$$\frac{d^2\phi_l(\lambda, r)}{\mathrm{d}r^2} + \left[k^2 - \lambda U(r) - \frac{l(l+1)}{r^2}\right]\phi_l(\lambda, r) = 0. \tag{14.33}$$

The equation with λ replaced by $(\lambda + \nabla\lambda)$ is

$$\frac{d^2\phi_l(\lambda+\nabla\lambda, r)}{\mathrm{d}r^2} + \left[k^2 - (\lambda+\nabla\lambda)U(r) - \frac{l(l+1)}{r^2}\right]\phi_l(\lambda+\nabla\lambda, r) = 0.$$

Take complex conjugate of the last equation

$$\frac{d^2\phi_l^*(\lambda+\nabla\lambda, r)}{\mathrm{d}r^2} + \left[k^2 - (\lambda+\nabla\lambda)U(r) - \frac{l(l+1)}{r^2}\right]\phi_l^*(\lambda+\nabla\lambda, r) = 0. \tag{14.34}$$

Multiply Eq. (14.33) by $\phi_l^*(\lambda+\nabla\lambda, r)$ and Eq. (14.34) by $\phi_l(\lambda, r)$ and subtract

$$\phi_l^*(\lambda+\nabla\lambda, r)\frac{d^2\phi_l(\lambda, r)}{\mathrm{d}r^2} - \phi_l(\lambda, r)\frac{d^2\phi_l^*(\lambda+\nabla\lambda, r)}{\mathrm{d}r^2} = -\nabla\lambda\, U(r)\phi_l^*(\lambda+\nabla\lambda, r)\phi_l(\lambda, r)$$

or

$$\frac{d}{\mathrm{d}r}\left[\phi_l^*(\lambda+\nabla\lambda, r)\frac{\mathrm{d}\phi_l(\lambda, r)}{\mathrm{d}r} - \phi_l(\lambda, r)\frac{\mathrm{d}\phi_l^*(\lambda+\nabla\lambda, r)}{\mathrm{d}r}\right] = -\nabla\lambda\, U(r)\phi_l^*(\lambda+\nabla\lambda, r)\phi_l(\lambda, r),$$

integrate over r from $r = 0$ to $r = R$. $(R > a)$, where a is the range of potential, so that at $r = R$, the wave function has attained its asymptotic form

$$\left[\phi_l^*(\lambda+\nabla\lambda, r)\frac{\mathrm{d}\phi_l(\lambda, r)}{\mathrm{d}r} - \phi_l(\lambda, r)\frac{\mathrm{d}\phi_l^*(\lambda+\nabla\lambda, r)}{\mathrm{d}r}\right]\Bigg|_{r=0}^{r=R} = -\nabla\lambda\int_0^R U(r)\phi_l^*(\lambda+\nabla\lambda, r)\phi_l(\lambda, r)\mathrm{d}r.$$

Now the asymptotic form (where $V(r) = 0$) is from Eq. (14.22)

$$\phi_l(\lambda, r) \underset{r\to\infty}{\longrightarrow} r R_l(\lambda, r) = \frac{C_l}{k}\sin\left(kr - \frac{l\pi}{2} + \delta_l(\lambda)\right).$$

Substituting it on the LHS of the previous equation, we get [note that $\phi_l(\lambda, r)|_{r=0}$ $= 0$ and $\phi_l(\lambda+\nabla\lambda, r)|_{r=0} = 0$. Also, since we are going to take the limit $\nabla\lambda \to 0$, we write C_l for both λ and $\lambda + \nabla\lambda$]

$$\frac{|C_l|^2}{k}\Big[\sin\Big(kR-\frac{l\pi}{2}+\delta_l(\lambda+\nabla\lambda)\Big)\cos\Big(kR-\frac{l\pi}{2}+\delta_l(\lambda)\Big)$$
$$-\sin\Big(kR-\frac{l\pi}{2}+\delta_l(\lambda)\Big)\cos\Big(kR-\frac{l\pi}{2}+\delta_l(\lambda+\nabla\lambda)\Big)\Big]$$
$$=-\nabla\lambda\int_0^R U(r)\phi_l^*(\lambda+\nabla\lambda,r)\phi_l(\lambda,r)\mathrm{d}r.$$

The square bracket of left side is

$$=\sin\Big((kR-\frac{l\pi}{2}+\delta_l(\lambda+\nabla\lambda))-(kR-\frac{l\pi}{2}+\delta_l(\lambda))\Big)$$
$$=\sin\Big(\delta_l(\lambda+\nabla\lambda)-\delta_l(\lambda)\Big).$$

Now take $\nabla\lambda$ small so that $\delta_l(\lambda+\nabla\lambda)-\delta_l(\lambda)$ is small and the above becomes

$$\sin\{\delta_l(\lambda+\nabla\lambda)-\delta_l(\lambda)\}\simeq\delta_l(\lambda+\nabla\lambda)-\delta_l(\lambda).$$

Hence

$$\frac{\delta_l(\lambda+\nabla\lambda)-\delta_l(\lambda)}{\nabla\lambda}=-\frac{k}{|C_l|^2}\int_0^R U(r)\phi_l^*(\lambda+\nabla\lambda,r)\phi_l(\lambda,r)\mathrm{d}r.$$

Take the limit $\nabla\lambda\to 0$

$$\lim_{\nabla\lambda\to 0}\Big[\frac{\delta_l(\lambda+\nabla\lambda)-\delta_l(\lambda)}{\nabla\lambda}\Big]=\lim_{\nabla\lambda\to 0}\Big(-\frac{k}{|C_l|^2}\int_0^R U(r)\phi_l^*(\lambda+\nabla\lambda,r)\phi_l(\lambda,r)\mathrm{d}r\Big).$$

Hence

$$\frac{\mathrm{d}\delta_l(\lambda)}{\mathrm{d}\lambda}=-\frac{k}{|C_l|^2}\int_0^R U(r)|\phi_l(\lambda,r)|^2\mathrm{d}r. \tag{14.35}$$

Now for $\lambda=0$ (no potential), there is no scattering. Hence there is no phase shift and $\delta_l(0)=0$. From Eq. (14.35), we see that $\frac{\mathrm{d}\delta_l(\lambda)}{\mathrm{d}\lambda}$ is negative. Then a plot of $\delta_l(\lambda)$ against λ for λ near zero has a typical form as shown in Fig. 14.5.
From this figure we can see that

for $\lambda>0$ (repulsive pot) $\delta_l(\lambda)<0$,
for $\lambda<0$ (attractive pot) $\delta_l(\lambda)>0$;

thus phase shift is positive for attractive potentials and negative for repulsive potentials. One can see this qualitatively as follows. When $V(r)=0$, i.e., $U(r)=0$ and the equation for $R_l^{(0)}(r)$ [superscript (0) indicates free particles] is

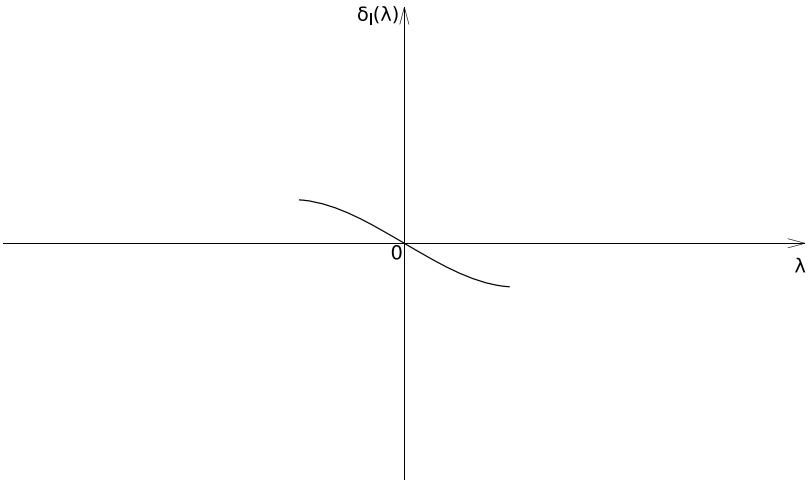


Fig. 14.5 Schematic plot of $\delta_l(\lambda)$ versus λ near $\lambda = 0$

$$\frac{d^2 R_l^{(0)}}{\mathrm{d}r^2} + \frac{2}{r}\frac{\mathrm{d}R_l^{(0)}}{\mathrm{d}r} + \left[k^2 - \frac{l(l+1)}{r^2}\right] R_l^{(0)} = 0.$$

This is spherical Bessel equation and the general solution is a linear combination of $j_l(kr)$ and $n_l(kr)$ (see Chap. 7, Sect. 7.2). Since $r = 0$ is included, we cannot retain $n_l(kr)$ (which diverges at $r = 0$) and we have

$$R_l^{(0)}(r) = A_l\, j_l(kr).$$

Consider for simplicity $l = 0$ only.

$$R_0^{(0)}(r) = A_0 j_0(kr) = A_0 \frac{\sin(kr)}{kr},$$

where $j_0(kr) = \frac{\sin(kr)}{kr}$ (Arfken 1966). Hence corresponding $\phi_0^{(0)}(r) = r R_0^{(0)}$ is

$$\phi_0^{(0)}(r) = C_0 \sin(kr) \qquad (C_0 = \frac{A_0}{k}).$$

In Fig. 14.6, $\phi_0^{(0)}(r)$ is shown by a dashed curve. It is a pure sine wave. Now for an attractive potential $V(r)$, the wave function of the particle ($\phi_0(r)$, represented by a continuous curve) will be *pulled inside the well*, since the particle has a higher probability to be in the region $r < a$, but once the particle is outside the range of potential (a), it does not feel any force and therefore its wave function is a sinusoidal wave again (which smoothly joins with the inside wave function) with wave length given by $\frac{2\pi}{k}$ [same as that of $\phi_0^{(0)}(r)$]. Hence a given phase (say $= 2\pi$) of $\phi_0(r)$

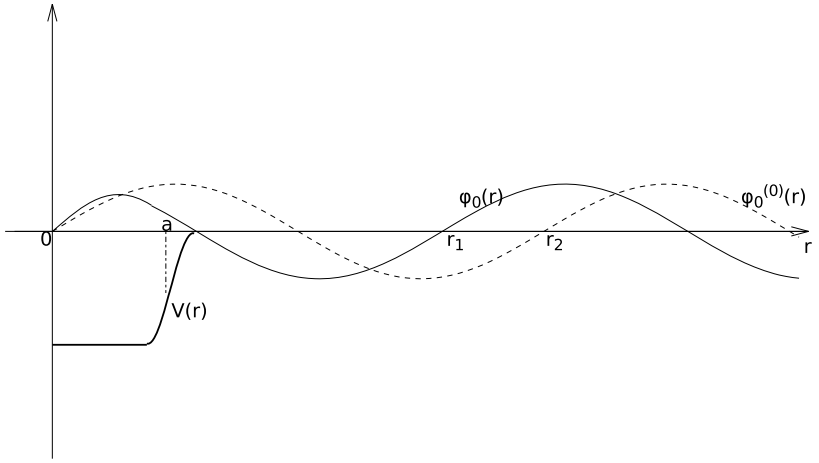


Fig. 14.6 Plot of $\phi_0(r)$ and $\phi_0^{(0)}(r)$ against r for an attractive potential $V(r)$. Range of the potential (a) is indicated by a vertical dash-dot line (note that without a specific definition, the *range* is only approximate and not well defined). $\phi_0(r)$ is pulled inside the well for the attractive $V(r)$

will occur at a smaller value of r (at $r = r_1$) than that for $\phi_0^{(0)}(r)$ (at $r = r_2$). The asymptotic form of $\phi_0(r)$ is

$$\phi_0(r)|_{r>a} \to C_0' \sin(kr + \delta_0) \qquad (C_0' = \text{ new constant}).$$

From Fig. 14.6, we see that

$$kr_1 + \delta_0 = 2\pi = kr_2. \qquad \text{Therefore} \quad \delta_0 = k(r_2 - r_1) > 0.$$

Hence phase shift is positive for an attractive potential.

The situation for a repulsive potential is shown in Fig. 14.7.
In this case, the wave $\phi_0(r)$ is *pushed out by the repulsive potential* and consequently a given phase (say $= 2\pi$) occurs at a larger value of $r = r_1$

$$kr_1 + \delta_0 = 2\pi = kr_2. \qquad \text{Therefore} \quad \delta_0 = k(r_2 - r_1) < 0.$$

Hence for a repulsive potential, the phase shift is negative.

These qualitative results agree with our earlier conclusions from Eq. (14.35) and Fig. 14.5.

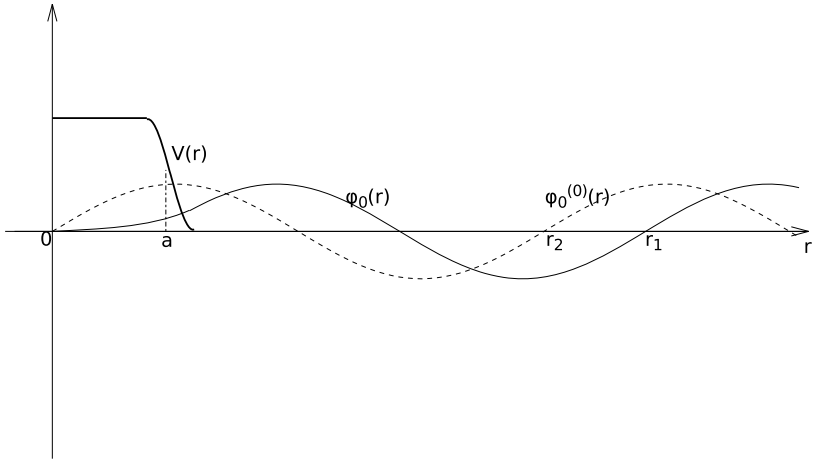


Fig. 14.7 Plot of $\phi_0(r)$ and $\phi_0^{(0)}(r)$ against r for a repulsive potential $V(r)$. Range of the potential (a) is indicated by a vertical dash-dot line (see note about range in Fig. 14.6). $\phi_0(r)$ is pushed out for the repulsive $V(r)$

14.4.4 Ramsauer–Townsend Effect

Consider an attractive potential (Fig. 14.8) which is strong enough, so that $\delta_0 = 180°$. Then from Eq. (14.27) the $l = 0$ partial cross-section is

$$\sigma_0 = \frac{4\pi}{k^2} \sin^2 \delta_0 = 0, \qquad \text{(since } \delta_0 = 180°\text{)}.$$

Now if energy is small enough, so that $ka << \sqrt{2}$ [see Eq. (14.29)], then only $l = 0$ contributes. Hence in this case $\sigma = \sigma_0 = 0$. This is called Ramsauer–Townsend (RT) effect.

For this to happen, the attractive potential $V(r)$ should be so strong that it pulls the $l = 0$ partial wave, $\phi_0(r)$ inside the well, such that its second node coincides with the first node of the $l = 0$ partial wave of the free particle having the same energy, $\phi_0^{(0)}(r)$. This gives $\delta_0 = 180°$. The energy of the incident particle should be small, so that only $l = 0$ partial wave is affected. This situation is shown in Fig. 14.8. Note that outside the range of $V(r)$, $\phi_0(r)$ has the same wave length as that of $\phi_0^{(0)}(r)$.

Thus, for a strong short-ranged attractive potential at a particular low energy the scattering cross-section may vanish (i.e., there is no scattering). The RT effect was seen in the scattering of electrons by rare gas atoms at about 0.7 eV bombarding energy. A rare gas atom consists entirely of closed shells and is relatively small in size. The combined Coulomb force of the nucleus and the atomic electrons exerted on an incident electron (corresponding potential is called screened Coulomb potential) is strong and sharply defined, so that a range of potential (a) can be defined. In this case, the foregoing conditions are satisfied at $E \simeq 0.7$ eV. This can occur only at a given

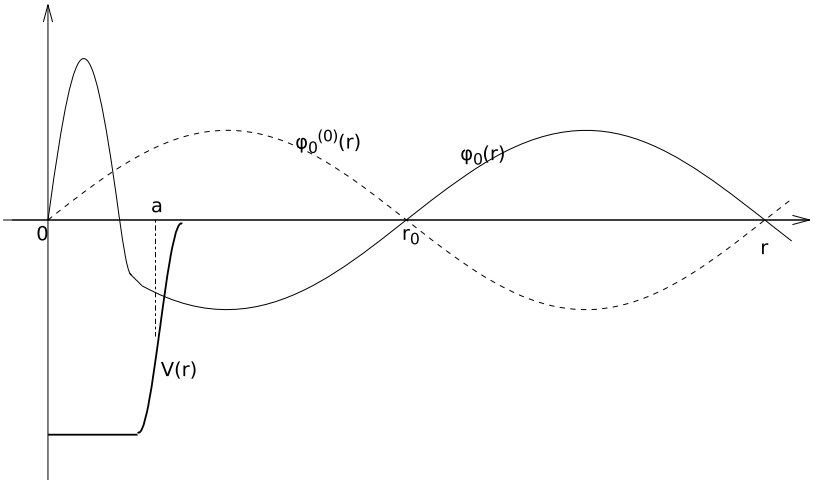


Fig. 14.8 Ramsauer–Townsend effect at a low energy (so that only $l = 0$ partial wave contributes) in an attractive potential. The phase shift δ_0 can be 180° if $\phi_0(r)$ is pulled in by the attractive potential such that its second node (at r_0) coincides with the first node of the free particle wave function $\phi_0^{(0)}(r)$

energy, since, although the interior wave function (for $r < a$) does not change much with small change of energy, (because energy is much smaller than magnitude of the depth of the attractive potential) the wave length $\frac{2\pi}{k}$ of the outside wave function (representing a free particle) changes rapidly and the condition $\delta_0 = 180^\circ$ can be satisfied only at a particular energy.

Physically, the RT effect may be thought of as a diffraction of electrons around the rare gas atoms, in which the wave function inside the atom is distorted in such a way that it fits smoothly to an undisturbed wave function outside. This is analogous to the perfect transmission at a particular energy through a one-dimensional rectangular barrier (see Sect. 13.2, Chap. 13), whenever the barrier contains an integral number of half wave lengths [see Eq. (13.6) and Fig. 13.3].

The RT effect is possible for an attractive potential only. In order that σ_0 is zero for a repulsive potential (for which phase shift is negative) we must have $\delta_0 = -180^\circ$. This is possible if the first node of $\phi_0(r)$ coincides with the second node (at r_0) of the free particle wave function $\phi_0^{(0)}(r)$ (see Fig. 14.9). Now outside the range of the potential, $\phi_0(r)$ becomes proportional to the free particle wave function with a negative phase shift. From Fig. 14.9, we see that for the shorter range potential $V_1(r)$ (with a range a_1, which is about a quarter of the free particle wave length), the outside wave function $(\phi_0)_1(r)$ becomes sinusoidal (with free particle wave length) and pushed out too soon to have its first node at r_0. The latter is possible only if the range of the potential is near the first node of $\phi_0^{(0)}(r)$, as seen for $V_2(r)$, whose range $a_2 \simeq \frac{\pi}{k}$. Thus for a repulsive potential δ_0 can be -180° only if $ka \simeq \pi \approx 3.12$. But since $ka > \sqrt{2}$, we see from Eq. 14.29 that $l = 1$ also contributes, for which δ_1 will

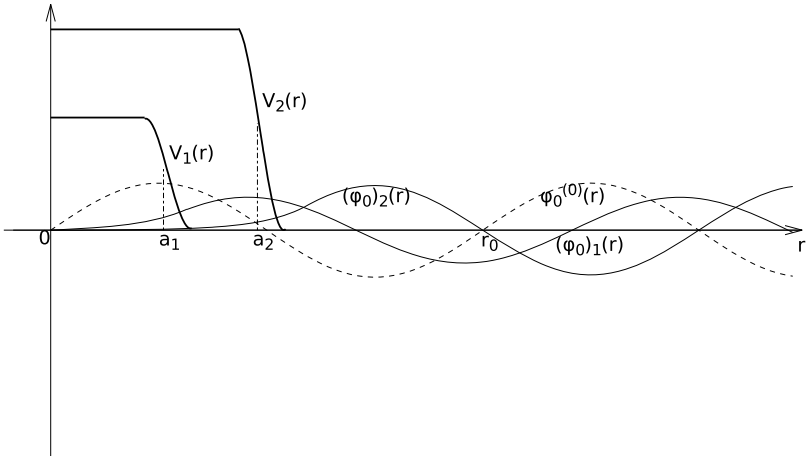


Fig. 14.9 For a repulsive potential there is no Ramsauer–Townsend effect. The phase shift δ_0 can be $-180°$ if $\phi_0(r)$ is pushed out by the repulsive potential such that its first node coincides with the second node of the free particle wave function $\phi_0^{(0)}(r)$ (dashed curve)

be different from $-180°$, and so $\sigma_1 \neq 0$. Hence σ can not vanish. Thus RT effect is not possible for a repulsive potential.

14.5 Scattering by a Perfectly Rigid Sphere

Consider the target to be a hard (i.e., perfectly rigid or impenetrable) sphere of radius a. Hence the potential is

$$\begin{aligned} V(r) &= +\infty \quad \text{for } r < a \\ &= 0 \qquad \text{for } r > a. \end{aligned} \tag{14.36}$$

Then clearly the wave function vanishes for $r < a$, (since a particle with any finite energy cannot penetrate the sphere). For $r > a$, the Schrödinger equation with $V(r) = 0$ is given by Eq. (14.19)

$$\frac{1}{r^2}\frac{d}{dr}(r^2\frac{dR_l}{dr}) + \left[k^2 - \frac{l(l+1)}{r^2}\right]R_l(r) = 0,$$

whose solution is [see Eq. (14.20)]

$$R_l(r) = C_l\Big[\cos\delta_l\, j_l(kr) - \sin\delta_l n_l(kr)\Big] \qquad (r > a).$$

At $r = a$, $R_l(a)$ must vanish

$$R_l(a) = C_l[\cos\delta_l j_l(ka) - \sin\delta_l n_l(ka)] = 0.$$

$$\text{Hence} \qquad \tan\delta_l = \frac{j_l(ka)}{n_l(ka)}. \tag{14.37}$$

Then the total cross-section from Eq. (14.27) is

$$\sigma = \frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l = \frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\frac{\tan^2\delta_l}{1+\tan^2\delta_l}$$
$$= \frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\frac{[j_l(ka)]^2}{[j_l(ka)]^2+[n_l(ka)]^2}. \tag{14.38}$$

We can study this in the low and high energy limits.

Low energy limit

For E very small, $ka << 1$, we can use the leading terms in the series expansion of $j_l(x)$ and $n_l(x)$ given in Ref. Arfken (1966)

$$j_l(x)\big|_{x\ \text{small}} \rightarrow \frac{x^l}{(2l+1)!!}$$
$$n_l(x)\big|_{x\ \text{small}} \rightarrow -\frac{(2l-1)!!}{x^{l+1}} \qquad (\text{for } l=0 \text{ it is } -\frac{1}{x}),$$

where $(2n)!! = 2.4.6.\cdots.(2n-2).2n$ and $(2n+1)!! = 1.3.5.\cdots.(2n-1).(2n+1)$, for $n =$ an integer ≥ 1, together with definitions $0!! = 1$ and $1!! = 1$. Substituting in the expression for σ, we get for low energy (only $l = 0$ contributes).

$$\sigma \underset{k\ \text{small}}{\longrightarrow} \frac{4\pi}{k^2}\frac{[j_0(ka)]^2}{[j_0(ka)]^2+[n_0(ka)]^2} = \frac{4\pi}{k^2}\sin^2(ka) \approx \frac{4\pi}{k^2}(ka)^2 = 4\pi a^2.$$

[The first equality in the above equation is an exact one, since $j_0(x) = \frac{\sin x}{x}$ and $n_0(x) = -\frac{\cos x}{x}$. Arfken (1966)]. This is four times the classical cross-section, which is the geometrical cross-section πa^2 (= area of the disk of radius a that blocks the propagation of the plane wave). This is not surprising, since at low energy, the wave length is very large and quantum effects are important. Hence we cannot expect the classical relation.

High energy limit

We may expect that at high energy, classical (geometrical) cross-section will result. At high energy many l-values contribute. The largest l-value is approximately given by Eq. (14.29)

$$l_{\max} = ka.$$

Hence

$$\sigma = \frac{4\pi}{k^2} \sum_{l=0}^{l_{\max}=ka} (2l+1)\sin^2\delta_l.$$

Now $$\sin^2\delta_l = \frac{\tan^2\delta_l}{1+\tan^2\delta_l} = \frac{[j_l(ka)]^2}{[j_l(ka)]^2+[n_l(ka)]^2}.$$

Since ka is large we can use asymptotic forms of $j_l(ka)$ and $n_l(ka)$, Eq. (14.21)

$$j_l(ka) \to \frac{1}{ka}\sin(ka - l\frac{\pi}{2}),$$
$$n_l(ka) \to -\frac{1}{ka}\cos(ka - l\frac{\pi}{2}).$$

Thus the phase shift δ_l decreases by $\frac{\pi}{2}$, each time l increases by 1. Hence

$$\begin{aligned}\sin^2\delta_l + \sin^2\delta_{l+1} &= \sin^2\delta_l + \sin^2(\delta_l - \frac{\pi}{2}) \\ &= \sin^2\delta_l + \cos^2\delta_l = 1.\end{aligned}$$

For large l, the factor $(2l+1)$ and $(2(l+1)+1)$ will not differ much. Hence it is legitimate to replace the sum of pairs of $\sin^2\delta_l$ for two consecutive l-values by its average value $\frac{1}{2}$. Then

$$\begin{aligned}\sigma \underset{ka\gg 1}{\longrightarrow} \frac{4\pi}{k^2}\sum_{l=0}^{ka}(2l+1).\frac{1}{2} &= \frac{2\pi}{k^2}\left[2\left\{\frac{ka(ka+1)}{2}\right\} + (ka+1)\right] \\ &= \frac{2\pi}{k^2}(ka+1)^2 \simeq 2\pi a^2.\end{aligned}$$

Thus the high energy scattering cross-section is twice the geometrical cross-section. This is due to the fact that there is a shadow scattering in addition to the normal scattering. We can write the scattering amplitude [see Eq. (14.25)]

$$\begin{aligned}f(\theta) &= \frac{1}{k}\sum_{l=0}^{ka}(2l+1)e^{i\delta_l}\sin\delta_l P_l(\cos\theta) \qquad \left(\text{Now } \sin\delta_l = \frac{e^{i\delta_l}-e^{-i\delta_l}}{2i}\right) \\ &= \frac{1}{2ik}\sum_{l=0}^{ka}(2l+1)e^{2i\delta_l}P_l(\cos\theta) + \frac{i}{2k}\sum_{l=0}^{ka}(2l+1)P_l(\cos\theta) \\ &\equiv f_{\text{reflection}}(\theta) + f_{\text{shadow}}(\theta).\end{aligned}$$

We can show [see Ref. Sakurai (2000)] that the contribution to σ from each of $|f_{\text{reflection}}|^2$ and $|f_{\text{shadow}}|^2$ is πa^2, while the cross term vanishes. Physically this is due to the following [see Ref. Schiff (1968)]. The hard sphere will produce a shadow in the forward direction of the plane wave (i.e., the parallel beam of incident particles). The cross-section of this is just πa^2, which is the geometrical cross-section. This is reflection scattering, as particles are scattered backwards. But particles with impact parameter larger than a will be scattered by the sphere into the shadow of the sphere in the forward direction (shadow scattering). Its value is also πa^2. Hence the total is $2\pi a^2$.

14.6 Coulomb Scattering

Consider a particle of charge $Z'e$ incident along $+\hat{z}$-axis on a target particle of charge Ze at the origin, interacting through the Coulomb potential. The relative Schrödinger equation (in c.g.s. units) is

$$\Big[-\frac{\hbar^2}{2\mu}\nabla_{\vec{r}}^2+\frac{ZZ'\mathrm{e}^2}{r}-E\Big]\psi(\vec{r})=0. \tag{14.39}$$

The Coulomb potential decreases as $\frac{1}{r}$ even in asymptotic region. Hence it is not a finite range potential and partial wave analysis is not advantageous. Besides spherical polar coordinates, the Schrödinger equation is separable also in parabolic coordinates. We follow Ref. Merzbacher (1965) and use parabolic coordinates (ξ, η, ϕ) (with $\hat{z}$-axis as the symmetry axis) defined in terms of Cartesian coordinates (x, y, z) as

$$\xi=r+z, \qquad \eta=r-z, \qquad \phi=\tan^{-1}\Big(\frac{y}{x}\Big). \tag{14.40}$$

In the parabolic coordinates, Eq. (14.39) becomes (see Ref. Arfken (1966) for $\nabla_{\vec{r}}^2$ in parabolic coordinates)

$$\begin{aligned}-\frac{\hbar^2}{2\mu}\Big[\frac{4}{\xi+\eta}\frac{\partial}{\partial\xi}\Big(\xi\frac{\partial}{\partial\xi}\Big)+\frac{4}{\xi+\eta}\frac{\partial}{\partial\eta}\Big(\eta\frac{\partial}{\partial\eta}\Big)+\frac{1}{\xi\eta}\frac{\partial^2}{\partial\phi^2}\Big]\psi(\xi,\eta,\phi)\\ +\Big[\frac{2ZZ'\mathrm{e}^2}{(\xi+\eta)}-E\Big]\psi(\xi,\eta,\phi)=0.\end{aligned} \tag{14.41}$$

Since the scattering process has axial symmetry, we look for solutions with axial symmetry, i.e., no ϕ dependence. Also the form of Eq. (14.41) suggests that ψ is separable in ξ and η. Hence we write

$$\psi(\xi,\eta,\phi)=f_1(\xi)f_2(\eta).$$

Substituting in Eq. (14.41) (and multiplying by $(-\frac{2\mu}{\hbar^2}.(\frac{\xi+\eta}{4}).\frac{1}{\psi})$), one can see that equations for f_1 and f_2 separate as

$$\frac{d}{\mathrm{d}\xi}\left(\xi\frac{\mathrm{d}f_1}{\mathrm{d}\xi}\right)+\frac{k^2}{4}\xi f_1-c_1f_1=0, \tag{14.42}$$

$$\frac{d}{\mathrm{d}\eta}\left(\eta\frac{\mathrm{d}f_2}{\mathrm{d}\eta}\right)+\frac{k^2}{4}\eta f_2-c_2f_2=0, \tag{14.43}$$

where $k^2=\frac{2\mu E}{\hbar^2}$, and c_1, c_2 are two constants of separation, subject to

$$c_1+c_2=\frac{ZZ'\mathrm{e}^2\mu}{\hbar^2}.$$

The asymptotic incident wave has the form

$$\mathrm{e}^{ikz}=\mathrm{e}^{i\frac{k}{2}(\xi-\eta)},$$

while the radially outgoing spherical wave has a factor

$$\mathrm{e}^{ikr}=\mathrm{e}^{i\frac{k}{2}(\xi+\eta)}.$$

Dependence on ξ of these functions guides us to choose a particular solution for $f_1(\xi)$ as

$$f_1(\xi)=\mathrm{e}^{i\frac{k}{2}\xi}.$$

We find that this in fact satisfies Eq. (14.42), with

$$c_1=i\frac{k}{2},$$

and hence is a particular solution. Then Eq. (14.43) becomes

$$\frac{d}{\mathrm{d}\eta}\left(\eta\frac{df_2}{\mathrm{d}\eta}\right)+\frac{k^2}{4}\eta f_2-\left(\frac{ZZ'\mathrm{e}^2\mu}{\hbar^2}-i\frac{k}{2}\right)f_2(\eta)=0.$$

We substitute

$$f_2(\eta)=\mathrm{e}^{-i\frac{k}{2}\eta}g(\eta).$$

This gives

$$\eta\frac{d^2g}{\mathrm{d}\eta^2}+(1-ik\eta)\frac{dg}{\mathrm{d}\eta}+n'kg(\eta)=0, \tag{14.44}$$

where

$$n' = -\frac{ZZ'e^2\mu}{\hbar^2 k} = -\frac{ZZ'e^2}{\hbar v}, \tag{14.45}$$

v being the relative speed, $\frac{\hbar k}{\mu}$. Substitute

$$\zeta = ik\eta$$

and multiply Eq. (14.44) by $\frac{1}{ik}$, to get

$$\zeta \frac{d^2 g(\zeta)}{d\zeta^2} + (1-\zeta)\frac{dg(\zeta)}{d\zeta} - in' g(\zeta) = 0. \tag{14.46}$$

Compare this equation with confluent hypergeometric equation [see Ref. Arfken (1966))]

$$xy''(x) + (c-x)y'(x) - ay(x) = 0, \tag{14.47}$$

whose first solution is $y(x) = {}_1F_1(a, c; x)$, called the confluent hypergeometric function, which is the first of two linearly independent solutions. This solution is regular at $x = 0$ for $c \neq 0, -1, -2, -3, \cdots$. Comparing Eqs. (14.46) and (14.47), we see that they have the same form, with $a = in'$, $c = 1$. Since the wave function must be regular at $\eta = 0$ (which corresponds to the forward direction, $r = z$), we choose the first solution, which is regular for $c = 1$. Choosing this solution of Eq. (14.46), we have an acceptable $g(\eta)$ as

$$g(\eta) = {}_1F_1(in', 1; ik\eta).$$

The asymptotic expansion (for large $|\eta|$, i.e., for large r, except in the forward direction, for which $r = z$) is [see Ref. Merzbacher (1965)]

$$g(\eta) \underset{|\eta| \text{ large}}{\longrightarrow} \frac{\exp(-n'\frac{\pi}{2})}{|\Gamma(1+in')|}\Big[\exp\Big\{-i(n' \ln k\eta - \sigma_{n'})\Big\} + \frac{n'}{k\eta}\exp\Big\{i(k\eta + n' \ln k\eta - \sigma_{n'})\Big\}\Big],$$

where $\Gamma(1+in') = |\Gamma(1+in')|e^{i\sigma_{n'}}$. Hence the complete solution is

$$\begin{aligned}\psi(\vec{r}) &= f_1(\xi) f_2(\eta) = e^{i\frac{k}{2}\xi} e^{-i\frac{k}{2}\eta} g(\eta) \\ &= e^{ikz} g(\eta) \\ &\underset{|r-z| \text{ large}}{\overset{\text{asymptotic}}{\longrightarrow}} \frac{\exp(-n'\frac{\pi}{2})}{|\Gamma(1+in')|}\Big[\exp\Big\{i\Big(kz - n' \ln k(r-z) + \sigma_{n'}\Big)\Big\} \\ &\quad + \frac{n'}{k(r-z)}\exp\Big\{i\Big(kr + n' \ln k(r-z) - \sigma_{n'}\Big)\Big\}\Big].\end{aligned} \tag{14.48}$$

Thus, for Coulomb scattering, the asymptotic form of the wave function is not $e^{ikz} + f(\theta)\frac{e^{ikr}}{r}$. This is because both the incident wave and the outgoing spherical wave are modified by the infinite range Coulomb force (which goes as $\frac{1}{r}$ and *not faster than* $\frac{1}{r}$), even when these are very far from the scatterer. We can verify that e^{ikz} is not a solution of Eq. (14.39) even for very large r. One can check that

$$e^{i\{kz - n' \ln k(r-z)\}} \tag{14.49}$$

is an asymptotic solution of Eq. (14.39) (to order $\frac{1}{r^2}$) for large r, (except in the forward direction, $r = z$ or $\eta = 0$). The wave function (14.49) is Coulomb modified plane wave. Similarly the Coulomb modified outgoing spherical wave is

$$\frac{e^{i\{kr + n' \ln 2kr\}}}{r}. \tag{14.50}$$

The additional terms in both the phases vary slowly with r (because of the logarithmic term) and we can still define

$$\frac{da}{d\Omega} = |f(\theta)|^2,$$

where $f(\theta)$ is the factor in front of the Coulomb modified outgoing spherical wave (14.50), in the full scattered wave function (see below). This is in analogy with the case of finite ranged potentials. In Eq. (14.48), put

$$k(r - z) = kr(1 - \cos\theta) = 2kr \sin^2(\theta/2),$$

and get

$$\psi(\vec{r}) \underset{|r-z| \text{large}}{\longrightarrow} \frac{\exp(-\frac{n'\pi}{2}) e^{i\sigma_{n'}}}{|\Gamma(1 + in')|} \Big[e^{i\{kz - n' \ln k(r-z)\}} + \frac{e^{i\{kr + n' \ln 2kr\}}}{r} \Big\{ \frac{n'}{k(1 - \cos\theta)} e^{in' \ln(\sin^2(\theta/2)) - 2i\sigma_{n'}} \Big\} \Big].$$

As discussed above, we identify Coulomb scattering amplitude as

$$f(\theta) = \frac{n'}{2k \sin^2(\theta/2)} e^{in' \ln(\sin^2(\theta/2)) - 2i\sigma_{n'}}.$$

Then

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f(\theta)|^2 = \frac{(n')^2}{4k^2\sin^4(\theta/2)} = \Big(-\frac{ZZ'\mathrm{e}^2\mu}{\hbar^2 k}\Big)^2 \frac{1}{4k^2\sin^4(\theta/2)}$$
$$= \frac{(ZZ'\mathrm{e}^2)^2}{16E^2}\frac{1}{\sin^4(\theta/2)}, \qquad \Big(\text{since } \frac{\hbar^2k^2}{2\mu} = E\Big). \tag{14.51}$$

This agrees with the classical Rutherford scattering cross-section.

14.7 Green's Function in Scattering Theory

We have seen that the scattering wave function asymptotically takes the form

$$\psi(\vec{r}\,) \underset{r\to\infty}{\longrightarrow} C\Big[\mathrm{e}^{ikz} + f(\theta)\frac{\mathrm{e}^{ikr}}{r}\Big], \tag{14.52}$$

in which incident plane wave is along $+\hat{z}$ direction and for a central potential scattering amplitude $f(\theta)$ is independent of ϕ. We will consider a general case in which the scattering potential $V(\vec{r}\,)$ is not necessarily central.

Figure 14.10 shows the direction of incidence as $\vec{k}_\alpha$ and the scattered wave moves along the direction $\vec{k}_\beta$ at an angle θ, which is the scattering angle. Let $\vec{k}_\alpha$ and $\vec{k}_\beta$ be the wave vectors of the incident plane wave $\psi_{\text{in}}(\vec{r}\,)$, and the scattered wave. The incident plane wave is

$$\psi_{\text{in}}(\vec{r}\,) = C\mathrm{e}^{i\vec{k}_\alpha\cdot\vec{r}} \equiv u_{\vec{k}_\alpha}(\vec{r}\,),$$

and its momentum is $\hbar\vec{k}_\alpha$. The outgoing spherical wave in the direction $\vec{k}_\beta$ is given by

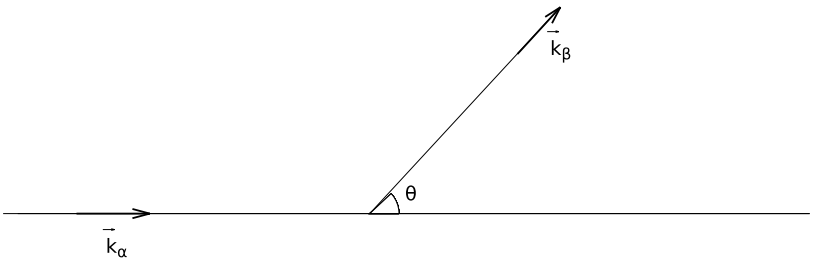


Fig. 14.10 Figure showing wave vectors $\vec{k}_\alpha$ and $\vec{k}_\beta$ for incident and scattered waves respectively and the scattering angle θ

$$\psi_{sc,\vec{k}_\beta}(\vec{r}\,) \propto \frac{e^{ik_\beta r}}{r}.$$

Then we rewrite Eq. (14.52) for the general case (note that for elastic scattering $|\vec{k}_\alpha| = |\vec{k}_\beta| = k$)

$$\begin{aligned}\psi_\alpha^{(+)}(\vec{r}\,) &\underset{r\to\infty}{\Longrightarrow} C\left[e^{i\vec{k}_\alpha\cdot\vec{r}} + f(\vec{k}_\alpha,\vec{k}_\beta)\frac{e^{ik_\beta r}}{r}\right] \\ &= C\left[e^{i\vec{k}_\alpha\cdot\vec{r}} + f(\vec{k}_\alpha,\vec{k}_\beta)\frac{e^{ikr}}{r}\right]. \end{aligned} \tag{14.53}$$

We write $\psi_\alpha^{(+)}(\vec{r}\,)$ instead of $\psi(\vec{r}\,)$ to indicate that Eq. (14.53) corresponds to the scattered wave function with spherically outgoing ('outgoing' is indicated by the superscript $^{(+)}$) wave in $\vec{k}_\beta$ direction, while the direction of incidence is $\vec{k}_\alpha$. The direction of incidence is indicated by the subscript α. Note that C is the normalization constant of the plane wave $e^{i\vec{k}_\alpha\cdot\vec{r}}$, and it is equal to $\frac{1}{L^{3/2}}$ for box normalization, $\frac{1}{(2\pi)^{3/2}}$ for δ-function normalization and $\frac{1}{\sqrt{v}}$ for unit flux normalization (as we discussed earlier in this chapter). It is also the normalization constant of the full wave function, since the scattered wave is very small (compared to the incident wave) in magnitude. Then Eq. (14.53) can be written for any r as

$$\psi_\alpha^{(+)}(\vec{r}\,) = \psi_{\text{in}}(\vec{r}\,) + C\psi_{\text{sc}}(\vec{r}\,) = u_{\vec{k}_\alpha}(\vec{r}\,) + C\psi_{\text{sc}}(\vec{r}\,), \tag{14.54}$$

where $\psi_{\text{in}}(\vec{r}\,) = u_{\vec{k}_\alpha}(\vec{r}\,) = Ce^{i\vec{k}_\alpha\cdot\vec{r}}$ and $\psi_{\text{sc}}(\vec{r}\,)$ is the scattered wave. The asymptotic form of $\psi_{\text{sc}}(\vec{r}\,)$ is $f(\vec{k}_\alpha,\vec{k}_\beta)\frac{e^{ikr}}{r}$. The Schrödinger equation is [see Eq. (14.17)]

$$\left[H_0 + V(\vec{r}\,) - E\right]\psi_\alpha^{(+)}(\vec{r}\,) = \left[H_0 + V(\vec{r}\,) - E\right]\left(u_{k_\alpha}(\vec{r}\,) + C\psi_{\text{sc}}(\vec{r}\,)\right) = 0.$$

In the above,

$$H_0 u_{k_\alpha}(\vec{r}\,) = -\frac{\hbar^2}{2\mu}\nabla^2 Ce^{i\vec{k}_\alpha\cdot\vec{r}} = \frac{\hbar^2k^2}{2\mu}u_{k_\alpha} = Eu_{\vec{k}_\alpha}.$$

Hence

$$(H_0 - E)\psi_{\text{sc}}(\vec{r}\,) = -\frac{1}{C}V(\vec{r}\,)\psi_\alpha^{(+)}(\vec{r}\,). \tag{14.55}$$

Assuming, for the moment, that $\psi_\alpha^{(+)}(\vec{r}\,)$ is known, Eq. (14.55) is an inhomogeneous equation for $\psi_{\text{sc}}(\vec{r}\,)$ and can be solved by using Green's function technique [Schiff (1968), Merzbacher 1965)]. In the following we review the use of Green's function [Arfken (1966)] for a formal solution of an inhomogeneous differential equation of the above type, as also its expansion in the complete set of eigen functions of the corresponding eigen value equation. Using these we will get a solution of Eq. (14.55).

Mathematical preliminary: Use of Green's function for solving inhomogeneous differential equations

For definition and use of Green's function we follow Ref. Arfken (1966). Consider the inhomogeneous second order differential equation in the interval $[a, b]$

$$\mathcal{L}(x)y(x) + \lambda\omega(x)y(x) = f(x), \tag{14.56}$$

where $\mathcal{L}(x)$ is a Hermitian second order differential operator in x in Stürm–Liouville form. The constant λ and the weight function $\omega(x)$ are given quantities. The solution $y(x)$ should satisfy specified homogeneous boundary conditions at the boundaries $x = a$ and $x = b$. Its formal solution is

$$y(x) = -\int_a^b G(x, x')f(x')\mathrm{d}x', \tag{14.57}$$

where $G(x, x')$, a function of two variables and called *Green's function* satisfies,

$$\Big[\mathcal{L}(x) + \lambda\omega(x)\Big]G(x, x') = -\delta(x - x'), \tag{14.58}$$

in which x' appears as a parameter. The Green's function $G(x, x')$, as a function of x, must satisfy the *same boundary conditions* as satisfied by $y(x)$ at $x = a$ and $x = b$. We can easily verify that Eq. (14.57) is a formal solution by substituting Eq. (14.57) in Eq. (14.56) and using Eq. (14.58), together with the boundary conditions on $G(x, x')$.

Eigen function expansion of Green's function

Suppose $\mathcal{L}(x)$ satisfies an eigen value equation

$$\mathcal{L}(x)U_n(x) + \lambda_n\,\omega(x)U_n(x) = 0, \tag{14.59}$$

where $U_n(x)$ is the eigen function corresponding to eigen value λ_n. The eigen function $U_n(x)$ satisfies the same boundary conditions as $y(x)$ at $x = a$ and $x = b$. These eigen functions form a complete orthonormal set. Then $G(x, x')$ as a function of x can be expanded in the complete orthonormal set of eigen functions $\{U_n(x)\}$ as

$$G(x, x') = -\sum_n \frac{U_n(x)U_n^*(x')}{\lambda - \lambda_n}. \tag{14.60}$$

We can verify this by substituting Eq. (14.60) in the left side of Eq. (14.58), using Eq. (14.59) and then the closure relation for the set $\{U_n(x)\}$.

We next come back to Eq. (14.55). Note that in our present case the one-dimensional variable x is to be replaced by the three-dimensional variable $\vec{r}$, with

consequent single eigen value index n is to be replaced by three eigen value indices. Comparing Eq. (14.55) with Eq. (14.56), we have

$$\mathcal{L}(x) \to H_0(\vec{r}\,) = -\frac{\hbar^2}{2\mu}\nabla^2, \quad \lambda = -E = -\frac{\hbar^2 k^2}{2\mu},$$
$$\omega(x) = 1, \quad y(x) \to \psi_{\rm sc}(\vec{r}\,) \text{ and } f(x) \to -\frac{1}{C}V(\vec{r}\,)\psi_\alpha^{(+)}(\vec{r}\,).$$

The eigen value equation satisfied by the three-dimensional operator H_0 is

$$H_0(\vec{r}\,)\,u_{\vec{k}'}(\vec{r}\,) = -\frac{\hbar^2}{2\mu}\nabla^2 u_{\vec{k}'}(\vec{r}\,) = \frac{\hbar^2 k'^2}{2\mu}u_{\vec{k}'}(\vec{r}\,), \qquad \text{replace } n \to \vec{k}'.$$
$$\text{Hence} \quad \lambda_n \to -\frac{\hbar^2 (k')^2}{2\mu}, \qquad U_n(x) \to u_{\vec{k}'}(\vec{r}\,).$$

Also we have to incorporate changes required from one dimension to three dimension. Furthermore eigen value $(-\frac{\hbar^2(k')^2}{2\mu})$ is a continuous function of k'. Hence we replace the sum over n by a three-dimensional integral over $\vec{k}'$, *viz.* $\sum_n \to \int d^3k'$. Therefore Eq. (14.60) translates to

$$G^{(+)}(\vec{r}, \vec{r}\,') = -\int \frac{u_{\vec{k}'}(\vec{r}\,)u^*_{\vec{k}'}(\vec{r}\,')}{[(-\frac{\hbar^2 k^2}{2\mu}) - (-\frac{\hbar^2(k')^2}{2\mu}) - i\epsilon']}d^3k'. \tag{14.61}$$

The extra term $-ie'$ in the denominator is included to take care of the singularity at $k' = k$. We will see later that a positive e' corresponds to an outgoing wave boundary condition, while a negative e$'$ corresponds to an incoming wave boundary condition. We retain the superscript $^{(+)}$ in G to indicate that we are solving Eq. (14.55) for the outgoing wave $\psi_\alpha^{(+)}$.

Next the equivalent of Eq. (14.57) gives a formal solution of Eq. (14.55) as

$$\psi_{\rm sc}(\vec{r}\,) = \frac{1}{C}\int G^{(+)}(\vec{r}, \vec{r}\,')V(\vec{r}\,')\psi_\alpha^{(+)}(\vec{r}\,')d^3r', \tag{14.62}$$

where $G^{(+)}(\vec{r}, \vec{r}\,')$ satisfies [equivalent to Eq. (14.58)]

$$(H_0 - E)G^{(+)}(\vec{r}, \vec{r}\,') = -\delta(\vec{r} - \vec{r'}), \tag{14.63}$$

subject to outgoing wave boundary conditions. Substitution of Eq. (14.62) in Eq. (14.54) gives

$$\psi_\alpha^{(+)}(\vec{r}\,) = u_{\vec{k}_\alpha}(\vec{r}\,) + \int G^{(+)}(\vec{r}, \vec{r}\,')V(\vec{r}\,')\psi_\alpha^{(+)}(\vec{r}\,')d^3r'. \tag{14.64}$$

This equation is called the Lippmann–Schwinger equation and is equivalent to the Schrödinger equation and is exact. It is an integral equation for the unknown wave function $\psi_\alpha^{(+)}(\vec{r})$, and it has the boundary condition (outgoing wave boundary condition) built in it [through $G^{(+)}(\vec{r},\vec{r}\,')$]. To solve it, one has to know Green's function $G^{(+)}(\vec{r},\vec{r}\,')$. To determine it, we use eigen function expansion Eq. (14.61).

$$
\begin{aligned}
G^{(+)}(\vec{r},\vec{r}\,') &= \frac{1}{(2\pi)^3}\int \frac{\mathrm{e}^{i\vec{k}'.(\vec{r}-\vec{r}')}}{\frac{\hbar^2k^2}{2\mu}-\frac{\hbar^2(k')^2}{2\mu}+i\epsilon'}\,d^3k' \\
&= \frac{2\mu}{\hbar^2(2\pi)^3}\int \frac{\mathrm{e}^{i\vec{k}'.\vec{\rho}}}{k^2-(k')^2+i\epsilon}\,d^3k' \quad (\text{where } \vec{\rho}=\vec{r}-\vec{r}',\ \ \epsilon=\frac{2\mu\epsilon'}{\hbar^2}) \\
&= \frac{2\mu}{\hbar^2(2\pi)^2}\int_0^\infty \frac{(k')^2\mathrm{d}k'}{k^2-(k')^2+i\epsilon}\int_{-1}^{1}\mathrm{e}^{ik'\rho\cos\theta''}d(\cos\theta'') \\
&\qquad (\text{where } \theta'' = \text{angle between } \vec{k}' \text{ and } \vec{\rho}) \\
&= \frac{\mu}{2\pi^2\hbar^2}\frac{2}{\rho}\int_0^\infty \frac{\sin(k'\rho)k'\mathrm{d}k'}{k^2-(k')^2+i\epsilon} \equiv \frac{\mu}{\pi^2\hbar^2\rho}I, \qquad (14.65)
\end{aligned}
$$

since $\int_{-1}^{1}\mathrm{e}^{ik'\rho x}\mathrm{d}x = \frac{2\sin(k'\rho)}{k'\rho}$.

Calculation of I

$$
\begin{aligned}
I &= -\int_0^\infty \frac{k'\sin(k'\rho)\mathrm{d}k'}{(k')^2-k^2-i\epsilon} = -\frac{1}{2}\int_{-\infty}^{\infty}\frac{k'\sin(k'\rho)\mathrm{d}k'}{(k')^2-k^2-i\epsilon}, \\
&= -\frac{1}{4i}\int_{-\infty}^{+\infty}\frac{k'(\mathrm{e}^{ik'\rho}-\mathrm{e}^{-ik'\rho})}{(k')^2-k^2-i\epsilon}\mathrm{d}k' \equiv -\frac{1}{4i}[I_1-I_2],
\end{aligned}
$$

$$
\text{where } I_1 = \int_{-\infty}^{+\infty}\frac{\mathrm{e}^{ik'\rho}k'\mathrm{d}k'}{(k')^2-k^2-i\epsilon} \quad \text{and} \quad I_2 = \int_{-\infty}^{+\infty}\frac{\mathrm{e}^{-ik'\rho}k'\mathrm{d}k'}{(k')^2-k^2-i\epsilon}.
$$

There are simple poles in both I_1 and I_2. The poles in I_1 and I_2 are at

$$
k' = \pm\sqrt{k^2+i\epsilon} = \pm k(1+\frac{i\epsilon}{k^2})^{\frac{1}{2}} \simeq \pm k(1+\frac{i\epsilon}{2k^2}) = \pm(k+\frac{i\epsilon}{2k}).
$$

We use a contour integration to evaluate I_1. Choose a contour c_1 as follows: along Re(k')-axis from $-R$ to R (where R is large), plus a semicircle (centered at the origin) of radius R in the upper-half k'-plane to complete the contour in the positive

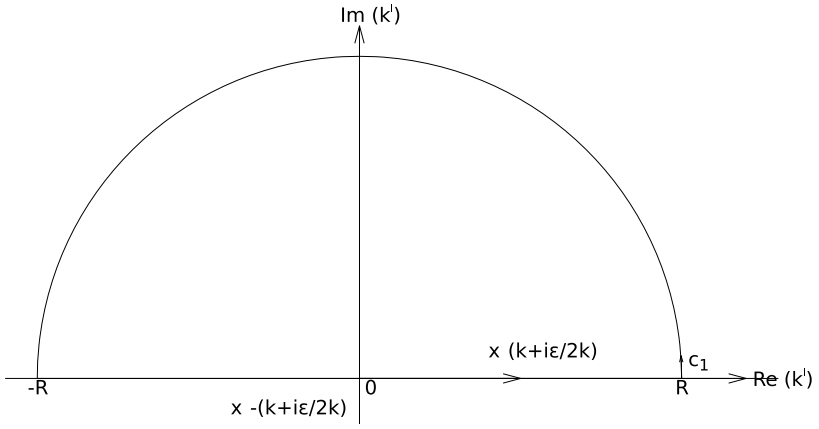


Fig. 14.11 Contour c_1 for I_1

(anticlockwise) sense. Only the pole at $k' = (k + \frac{i\epsilon}{2k})$ is enclosed (see Fig. 14.11, we take $\epsilon > 0$). Finally we take the limit $R \to \infty$.

We can show that the contribution of the semicircle (on which $k' = Re^{i\vartheta}$) vanishes in the limit $R \to \infty$, since $e^{ik'\rho} = e^{-R\sin\vartheta}.e^{iR\cos\vartheta} \to 0$ (in the upper-half plane, where $0 < \vartheta < \pi$) for $R \to \infty$. Hence

$$\begin{aligned} I_1 &= \lim_{R\to\infty} \oint_{c_1} = 2\pi i \times \text{residue of pole enclosed in } c_1 \\ &= 2\pi i \times \Big(k' - (k + \frac{i\epsilon}{2k})\Big) \frac{e^{ik'\rho}\, k'}{(k')^2 - (k + \frac{i\epsilon}{2k})^2}\Big|_{k'=k+\frac{i\epsilon}{2k}} \\ &= \pi i\, e^{i(k+\frac{i\epsilon}{2k})\rho} \underset{\epsilon\to 0}{\longrightarrow} \pi i\, e^{ik\rho}. \end{aligned} \tag{14.66}$$

For the second integral (which has $e^{-ik'\rho}$) we must close the contour c_2 in the lower half plane (so that the contribution from the semicircle vanishes). In this case the sense of the contour is clockwise, i.e., negative (see Fig. 14.12). Hence only the simple pole at $k' = -(k + \frac{i\epsilon}{2k})$ is enclosed and we have

$$\begin{aligned} I_2 &= \lim_{R\to\infty} \oint_{c_2} \\ &= -2\pi i \times \Big(k' + (k + \frac{i\epsilon}{k^2})\Big) \frac{e^{-ik'\rho}\, k'}{(k')^2 - (k + \frac{i\epsilon}{2k})^2}\Big|_{k'=-(k+\frac{i\epsilon}{2k})} \\ &= -\pi i\, e^{+i(k+\frac{i\epsilon}{2k})\rho} \underset{\epsilon\to 0}{\longrightarrow} -\pi i\, e^{ik\rho}, \end{aligned}$$

where the negative sign in the second equality is due to clockwise sense of c_2. Therefore, including contribution from Eq. (14.66), we have

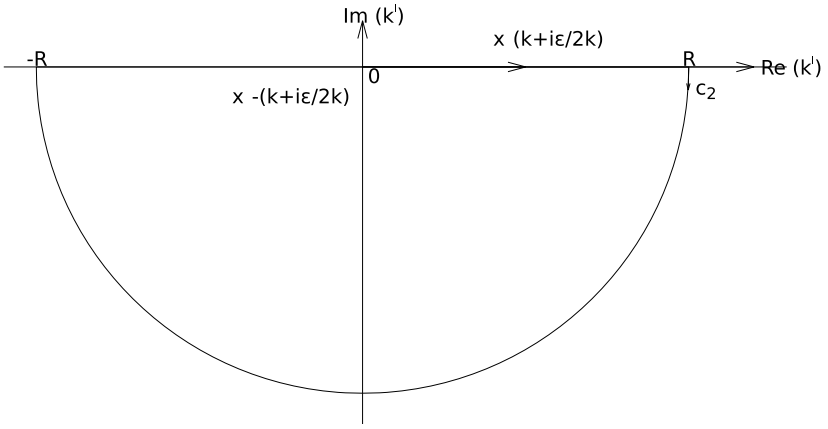


Fig. 14.12 Contour c_2 for I_2

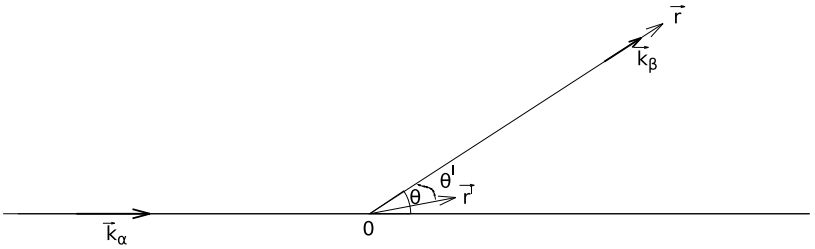


Fig. 14.13 Schematic display of $\vec{k}_\alpha, \vec{k}_\beta, \vec{r}, \vec{r}\,'$ and θ'

$$I = -\frac{1}{4i}[I_1 - I_2] = -\frac{1}{4i}[\pi i\, \mathrm{e}^{ik\rho} - (-\pi i\, \mathrm{e}^{ik\rho})] = -\frac{\pi}{2}\mathrm{e}^{ik\rho}. \tag{14.67}$$

If we took $\epsilon < 0$, poles would be reflected by the real axis and we would get $I = -\frac{\pi}{2}\,\mathrm{e}^{-ik\rho}$, which corresponds to an incoming wave. Thus we see that the sign of ϵ determines the asymptotic nature of the waves and hence the boundary condition to be imposed [Arfken (1966), Merzbacher (1965)].

Substituting Eq. (14.67) in Eq. (14.65), we have Green's function

$$\begin{aligned} G^{(+)}(\vec{r}, \vec{r}\,') &= \frac{\mu}{\pi^2\hbar^2\rho}(-\frac{\pi}{2}\mathrm{e}^{ik\rho}) = -\frac{\mu}{2\pi\hbar^2}\frac{\mathrm{e}^{ik\rho}}{\rho} \\ &= -\frac{\mu}{2\pi\hbar^2}\frac{\mathrm{e}^{ik|\vec{r}-\vec{r}\,'|}}{|\vec{r}-\vec{r}\,'|}. \end{aligned} \tag{14.68}$$

Then from Eq. (14.62)

$$\psi_{\mathrm{sc}}(\vec{r}\,) = -\frac{\mu}{2\pi\hbar^2\, C}\int \frac{\mathrm{e}^{ik|\vec{r}-\vec{r}\,'|}}{|\vec{r}-\vec{r}\,'|} V(\vec{r}\,')\psi_\alpha^{(+)}(\vec{r}\,')d^3r'. \tag{14.69}$$

Now for very large r (asymptotic region), the variable of integration $r' = |\vec{r}\,'|$ must be small, since $V(\vec{r}\,')$ has a finite range. Then for r very large and r' finite, we have, retaining terms of order $\mathcal{O}(\frac{r'}{r})$ (see Fig. 14.13)

$$\begin{aligned}|\vec{r} - \vec{r}\,'| &= \left[r^2 + (r')^2 - 2rr'\cos\theta'\right]^{\frac{1}{2}} \\ &\approx r\left[1 - 2\frac{r'}{r}\cos\theta'\right]^{\frac{1}{2}} \\ &\approx r(1 - \frac{r'}{r}\cos\theta') = r - r'\cos\theta'.\end{aligned}$$

where θ' is the angle between $\vec{r}$ and $\vec{r}\,'$. Also

$$\begin{aligned}\frac{1}{|\vec{r} - \vec{r'}|} &= \frac{1}{[r^2 + (r')^2 - 2rr'\cos\theta']^{\frac{1}{2}}} \approx \frac{1}{r}[1 - 2\frac{r'}{r}\cos\theta']^{-\frac{1}{2}} \\ &\approx \frac{1}{r}(1 + \frac{r'}{r}\cos\theta') = \frac{1}{r} + \frac{r'}{r^2}\cos\theta' \approx \frac{1}{r},\end{aligned}$$

neglecting small term $\frac{r'}{r^2}$. Then from Eq. (14.69)

$$\begin{aligned}\psi_{sc}(\vec{r}) &\xrightarrow[r\ \text{large}]{} -\frac{\mu}{2\pi\hbar^2 C}\int \frac{e^{ik(r - r'\cos\theta')}}{r} V(\vec{r}\,')\psi_\alpha^{(+)}(\vec{r}\,')d^3r' \\ &= \frac{e^{ikr}}{r}\Big[-\frac{\mu}{2\pi\hbar^2 C}\int e^{-ikr'\cos\theta'} V(\vec{r}\,')\psi_\alpha^{(+)}(\vec{r}\,')d^3r'\Big] \\ &\qquad\qquad \text{Note that } kr'\cos\theta' = |\vec{k}_\beta|r'\cos\theta' = \vec{k}_\beta.\vec{r}\,' \\ &= \frac{e^{ikr}}{r}\Big[-\frac{\mu}{2\pi\hbar^2 C}\int e^{-i\vec{k}_\beta.\vec{r}\,'} V(\vec{r}\,')\psi_\alpha^{(+)}(\vec{r}\,')d^3r'\Big] \\ &= \frac{e^{ikr}}{r}\Big[-\frac{\mu}{2\pi\hbar^2\,|C|^2}\int u^*_{\vec{k}_\beta}(\vec{r}\,')V(\vec{r}\,')\psi_\alpha^{(+)}(\vec{r}\,')d^3r'\Big],\end{aligned}$$

where $u_{\vec{k}_\beta}(\vec{r}) = Ce^{i\vec{k}_\beta\cdot\vec{r}}$ is the plane wave in the $\vec{k}_\beta$-direction. Comparing $\psi_{sc}(\vec{r})$ with $f(\vec{k}_\alpha, \vec{k}_\beta)\frac{e^{ikr}}{r}$, we get

$$f(\vec{k}_\alpha, \vec{k}_\beta) = -\frac{\mu}{2\pi\hbar^2\,|C|^2}\int u^*_{\vec{k}_\beta}(\vec{r}\,')V(\vec{r}\,')\psi_\alpha^{(+)}(\vec{r}\,')d^3r'. \tag{14.70}$$

This is a formal relation, since we do not know $\psi_\alpha^{(+)}(\vec{r}\,)$ yet, which must be obtained by solving the Lippmann–Schwinger equation (14.64). In the following we discuss an approximation method to achieve that goal.

14.8 Born Approximation

Equation (14.70) is not convenient for calculating $f(\vec{k}_\alpha, \vec{k}_\beta)$, since $\psi_\alpha^{(+)}(\vec{r}\,')$ is not known. The latter can be obtained by solving the Lippmann–Schwinger equation

$$\psi_\alpha^{(+)}(\vec{r}\,) = u_{\vec{k}_\alpha}(\vec{r}\,) + \int G^{(+)}(\vec{r}, \vec{r}\,')V(\vec{r}\,')\psi_\alpha^{(+)}(\vec{r}\,')d^3r'.$$

Magnitude of the second term of the right side is much smaller than that of the first term, since magnitude of the scattered wave is much smaller than that of the incident wave [see Eq. (14.54)]. Hence this integral equation can be solved as a series obtained by substituting for $\psi_\alpha^{(+)}$ within the integral, from this equation itself, successively

$$\begin{aligned}
\psi_\alpha^{(+)}(\vec{r}) &= u_{\vec{k}_\alpha}(\vec{r}) + \int G^{(+)}(\vec{r}, \vec{r}\,')V(\vec{r}\,')\Big[u_{\vec{k}_\alpha}(\vec{r}\,') \\
&\quad + \int G^{(+)}(\vec{r}\,', \vec{r}\,'')V(\vec{r}\,'')\psi_\alpha^{(+)}(\vec{r}\,'')d^3r''\Big]d^3r' \\
&= u_{\vec{k}_\alpha}(\vec{r}) + \int G^{(+)}(\vec{r}, \vec{r}\,')V(\vec{r}\,')u_{\vec{k}_\alpha}(\vec{r}\,')d^3r' \\
&\quad + \int G^{(+)}(\vec{r}, \vec{r}\,')V(\vec{r}\,')\Big(\int G^{(+)}(\vec{r}\,', \vec{r}\,'')V(\vec{r}\,'')u_{\vec{k}_\alpha}(\vec{r}\,'')d^3r''\Big)d^3r' \\
&\quad + \cdots . \qquad (14.71)
\end{aligned}$$

This is called the "Born Series". From the above argument, the series is likely to converge rapidly. Terminating the series after n terms leads to the nth Born approximation. Since $u_{\vec{k}_\alpha}(\vec{r}\,) = Ce^{i\vec{k}_\alpha\cdot\vec{r}}$ and Green's function [see Eq. (14.68)] are known, one can, in principle, obtain $\psi_\alpha^{(+)}(\vec{r}\,)$ as a series. The first Born approximation (or simply 'Born approximation') [Schiff (1968)] for calculating scattering amplitude $f(\vec{k}_\alpha, \vec{k}_\beta)$ consists of replacing $\psi_\alpha^{(+)}(\vec{r}\,')$ by $u_{\vec{k}_\alpha}(\vec{r}\,')$ on the right side of Eq. (14.70), [corresponding to the first term of the series (14.71)]. Then

$$\begin{aligned}
f(\vec{k}_\alpha, \vec{k}_\beta) \equiv f_B(\vec{k}_\alpha, \vec{k}_\beta) &\simeq -\frac{\mu}{2\pi\hbar^2\,|C|^2}\int u_{\vec{k}_\beta}^*(\vec{r}\,')V(\vec{r}\,')u_{\vec{k}_\alpha}(\vec{r}\,')d^3r' \\
&= -\frac{\mu}{2\pi\hbar^2}\int e^{i(\vec{k}_\alpha-\vec{k}_\beta)\cdot\vec{r}\,'}V(\vec{r}\,')d^3r' \qquad (\text{since } u_{\vec{k}_\alpha}(\vec{r}\,) = Ce^{i\vec{k}_\alpha\cdot\vec{r}}) \\
&= -\frac{\mu}{2\pi\hbar^2}\int V(\vec{r}\,')e^{i\vec{q}\cdot\vec{r}\,'}d^3r', \qquad (14.72)
\end{aligned}$$

where $\vec{q} = \vec{k}_\alpha - \vec{k}_\beta$. Hence $\hbar\vec{q}$ is the momentum transfer from the incident particle to the scattering potential, during the collision. The quantity $f_B(\vec{k}_\alpha, \vec{k}_\beta)$ is called the first Born scattering amplitude. Equation (14.72) shows that it is proportional to the Fourier transform [Arfken (1966)] of the potential as a function of the momentum transfer.

For a central potential $V(\vec{r}\,') = V(r')$, we can do the integrations over angles in Eq. (14.72). In this case f_B depends only on the scattering angle θ, not on ϕ.

Choose $\vec{q}$ as the z direction for $\vec{r}\,'$-integration. The polar coordinates of $\vec{r}\,'$ are $\vec{r}\,' = (r', \vartheta', \varphi')$. Then the integral in Eq. (14.72) is

$$\int_0^\infty V(r')\mathrm{e}^{iqr'\cos\vartheta'}(r')^2\mathrm{d}r'\sin\vartheta'\mathrm{d}\vartheta'\mathrm{d}\varphi'$$

$$= 2\pi\int_0^\infty V(r')(r')^2\mathrm{d}r'\int_{-1}^{1}\mathrm{e}^{iqr'\cos\vartheta'}d(\cos\vartheta') = \frac{4\pi}{q}\int_0^\infty V(r')r'\sin(qr')\mathrm{d}r',$$

Thus we have

$$f_B(\theta) = -\frac{2\mu}{\hbar^2 q}\int_0^\infty V(r')r'\sin(qr')\mathrm{d}r'. \tag{14.73}$$

where

$$q^2 = |\vec{k}_\alpha - \vec{k}_\beta|^2 = k^2 + k^2 - 2k^2\cos\theta \quad (\text{Because} \quad |\vec{k}_\alpha| = |\vec{k}_\beta| = k)$$
$$= 2k^2(1-\cos\theta) = 4k^2\sin^2\frac{\theta}{2}.$$

Therefore $q = 2k\sin\dfrac{\theta}{2}$. (14.74)

As an application of Born approximation, consider the Yukawa potential

$$V(r) = V_0\frac{\mathrm{e}^{-\varsigma r}}{\varsigma r}, \tag{14.75}$$

where V_0 is the strength and ς is the reciprocal of range of the potential. This potential comes as a simple theoretical prediction for nucleon-nucleon potential arising from the one-pion-exchange process. This is also the form of the screened Coulomb potential, which is the potential as seen by an incident electron approaching toward a neutral atom of nuclear change $+Ze$ surrounded by a negatively charged cloud produced by Z electrons. In this case $\frac{1}{\varsigma} = a$ is the "radius" of the atom and $V_0 a = -Z\mathrm{e}^2$ [so that the potential is $-Z\mathrm{e}^2\frac{\mathrm{e}^{-r/a}}{r}$]. On the other hand taking the limit $V_o \to 0, \varsigma \to 0$, subject to $\frac{V_0}{\varsigma} = ZZ'\mathrm{e}^2 = \text{finite}$, Eq. (14.75) gives the Coulomb potential:

$$\lim_{V_0\to 0,\varsigma\to 0,(V_0/\varsigma = ZZ'\mathrm{e}^2)} V(r) = \lim_{V_0\to 0,\varsigma\to 0,(V_0/\varsigma = ZZ'\mathrm{e}^2)} \frac{1}{r}\left(\frac{V_0}{\varsigma}\right)\left[1 - \varsigma r + \frac{\varsigma^2 r^2}{2} + \cdots\right]$$
$$= \frac{ZZ'\mathrm{e}^2}{r} \qquad \to \text{Coulomb potential.}$$

For the potential (14.75), using Eq. (14.73) we have the Born amplitude

$$\begin{aligned} f_B(\theta) &= -\frac{2\mu}{\hbar^2 q}.\frac{V_0}{\varsigma}\int_0^\infty \frac{\mathrm{e}^{-\varsigma r'}}{r'} r' \sin(qr')\mathrm{d}r' \\ &= -\frac{2\mu V_0}{\hbar^2 q\,\varsigma}\int_0^\infty \mathrm{e}^{-\varsigma r'}\sin(qr')\mathrm{d}r' \\ &= -\frac{2\mu V_0}{\hbar^2 q\,\varsigma}\left(\frac{q}{\varsigma^2+q^2}\right) \\ &= -\frac{2\mu}{\hbar^2}\left(\frac{V_0}{\varsigma}\right)\frac{1}{[\varsigma^2+4k^2\sin^2\theta/2]}, \end{aligned} \tag{14.76}$$

using Eq. (14.74). Therefore the differential scattering cross-section in first Born approximation is

$$\frac{\mathrm{d}\sigma_B}{\mathrm{d}\Omega} = |f_B(\theta)|^2 = \frac{4\mu^2}{\hbar^4}\left(\frac{V_0}{\varsigma}\right)^2\frac{1}{[\varsigma^2+4k^2\sin^2\theta/2]^2}.$$

The total cross-section is [Schiff (1968)]

$$\sigma_B = 2\pi\int_0^\pi |f_B(\theta)|^2\sin\theta\mathrm{d}\theta = \frac{16\mu^2}{\hbar^4}\pi\left(\frac{V_0}{\varsigma}\right)^2\frac{1}{\varsigma^2}\frac{1}{(4k^2+\varsigma^2)}. \tag{14.77}$$

We can easily obtain the Coulomb scattering result from Eq. (14.76) by taking the limit $V_0 \to 0,\ \varsigma \to 0$, subject to $\frac{V_0}{\varsigma} = ZZ'\mathrm{e}^2$

$$\begin{aligned} f_B(\theta)\Big|_{\text{Coulomb}} &= \lim_{V_0\to 0,\varsigma\to 0,(\frac{V_0}{\varsigma}=ZZ'e^2)} f_B(\theta) \\ &= -\frac{2\mu}{\hbar^2}ZZ'e^2\frac{1}{4k^2\sin^2\theta/2} = -\frac{ZZ'e^2}{4E}\frac{1}{\sin^2\theta/2}. \end{aligned}$$

$$\text{Therefore } \left(\frac{\mathrm{d}\sigma_B}{\mathrm{d}\Omega}\right)_{\text{Coulomb}} = \left|f_B(\theta)|_{\text{Coulomb}}\right|^2 = \frac{(ZZ'e^2)^2}{16E^2}\frac{1}{\sin^4\theta/2}.$$

This is in complete agreement with the exact result Eq. (14.51), although it has been obtained by Born approximation! This is one of the peculiarities shown by the Coulomb potential.

Validity of Born approximation.

Since the Born approximation consists of replacing $\psi_\alpha^{(+)}(\vec{r}\,')$ by $u_{k_\alpha}(\vec{r}\,')$ in the expression of $f(\vec{k}_\alpha, \vec{k}_\beta)$, requirement of its validity demands that $|\psi_{\text{sc}}|$ be small compared to $|\mathrm{e}^{i\vec{k}_\alpha.\vec{r}}| = 1$, since from Eq. (14.54) we have

$$\psi_\alpha^{(+)}(\vec{r}\,) = C[e^{i\vec{k}_\alpha.\vec{r}} + \psi_{sc}(\vec{r}\,)].$$

We require this at $r = 0$, where $|\psi_{sc}(\vec{r}\,)|$ is expected to be the largest. Then from Eq. (14.62)

$$\left|\psi_{sc}(0)\right| = \frac{1}{C}\left|\int G^{(+)}(0,\vec{r}\,')V(\vec{r}\,')\psi_\alpha^{(+)}(\vec{r}\,')d^3r'\right| \ll 1.$$

Replacing $\psi_\alpha^{(+)}(\vec{r}\,')$ by $u_{k_\alpha}(\vec{r}\,')$, $\quad \frac{1}{C}\left|\int G^{(+)}(0,\vec{r}\,')V(\vec{r}\,')u_{\vec{k}_\alpha}(\vec{r}\,')d^3r'\right| \ll 1.$

Using Eq. (14.68), we have for a central potential, $V(\vec{r}\,') = V(r')$

$$\frac{\mu}{2\pi\hbar^2}\left|\int \frac{e^{ikr'}}{r'}V(r')\int_{-1}^{1} e^{ikr'\cos\vartheta'}(r')^2 dr' d(\cos\vartheta')2\pi\right| \ll 1.$$

Doing the angular integration, we have the criterion for validity of Born approximation

$$\frac{\mu}{\hbar^2 k}\left|\int_0^\infty V(r')\left[e^{2ikr'} - 1\right]dr'\right| \ll 1,$$

$$\text{or} \quad \frac{2\mu}{\hbar^2 k}\left|\int_0^\infty V(r')e^{ikr'}\sin(kr')dr'\right| \ll 1. \tag{14.78}$$

We check the possibility of this criterion being satisfied at high and low energy limits.

High energy – At high energy, k is large, so that $e^{ikr'}\sin(kr')$ is a rapidly oscillating function of r'. Since $V(r')$ is a smoothly varying function of r', the integral becomes very small (due to cancelation in pairs of consecutive half-waves) for large k. In addition the $\frac{1}{k}$ factor also decreases. Hence at high energy, the criterion is automatically satisfied. Hence Born approximation is "good" at high energy.

Low energy – We have a general result

$$\left|\int_0^\infty V(r')e^{ikr'}\sin(kr')dr'\right| \le \int_0^\infty \left|V(r')e^{ikr'}\sin(kr')\right|dr'.$$

(This is true since the modulus of a sum is less than or equal to the sum of moduli of individual terms). Now at very low energy ($k \to 0$) and for a finite range potential (only finite values of r' contribute to the integral), $kr' < \pi$ and $\sin(kr')$ remains positive. Hence

$$\left|\int_0^\infty V(r')e^{ikr'}\sin(kr')dr'\right| \le \int_0^\infty \left|V(r')\right|\sin(kr')dr' \tag{14.79}$$

$$\xrightarrow[k \text{ very small}]{} k\int_0^\infty \left|V(r')\right|r'dr'.$$

Substituting this in Eq. (14.78), we have the validity criterion for Born approximation in the low energy limit

$$\frac{2\mu}{\hbar^2}\int_0^\infty |V(r')|r'dr' \ll 1$$

$$\text{or} \quad \int_0^\infty |V(r')|r'dr' \ll \frac{\hbar^2}{2\mu} \tag{14.80}$$

This means that the potential $V(r)$ should be so weak that the corresponding attractive well (by changing sign of $V(r)$ if necessary) should not support even one bound state (see below for the Yukawa potential; also see Ref. Sakurai (2000).

Application to Yukawa potential

For the Yuhawa potential $V(r) = V_0\frac{e^{-\varsigma r}}{\varsigma r}$, We have from Eq. (14.78) the condition of validity is

$$\frac{\mu|V_0|}{\hbar^2 k\,\varsigma}\left|\int_0^\infty (e^{2ikr'}-1)\frac{e^{-\varsigma r'}}{r'}dr'\right| \ll 1.$$

Evaluating the integral we have [see Ref. Merzbacher (1965)],

$$\frac{\mu|V_0|}{\hbar^2 k\,\varsigma}\sqrt{\left(\ln\sqrt{1+\frac{4k^2}{\varsigma^2}}\right)^2+\left(\tan^{-1}\frac{2k}{\varsigma}\right)^2} \ll 1. \tag{14.81}$$

Low energy: For low energy, $\frac{k}{\varsigma} << 1$, this simplifies to (since the second term is larger)

$$\frac{2\mu|V_0|}{\hbar^2\varsigma^2} \ll 1 \quad \rightarrow \text{independent of velocity.} \tag{14.82}$$

We get the same result as in Eq. (14.80). Using the range of potential $a = \frac{1}{\varsigma}$,

$$\frac{2\mu}{\hbar^2}|V_0|a^2 \ll 1 \quad \Rightarrow \quad |V_0|a^2 \ll \frac{\hbar^2}{2\mu}.$$

This means that in the low energy limit, the Born approximation is valid if the corresponding attractive potential is too weak to support a bound state. This agrees with our general conclusion obtained earlier.

For the limiting case of Yukawa potential going over to the Coulomb potential, $V_0 \to 0,\ \varsigma \to 0$, subject to $\frac{V_0}{\varsigma}$ constant and $= ZZ'e^2$. Then Eq. (14.82) becomes

$$\lim_{\varsigma \to 0} \left(\frac{2\mu Z Z' e^2}{\hbar^2 \varsigma} \right) \ll 1,$$

which cannot be satisfied. This is also true for screened Coulomb potential where $\frac{V_0}{\varsigma} = -Ze^2$ and $\frac{1}{\varsigma} = a =$ radius of the atom. Hence Eq. (14.82) becomes

$$\frac{2\mu Z e^2}{\hbar^2} a \ll 1.$$

For a moderately heavy atom $a \simeq \frac{\hbar^2}{\mu e^2 Z^{\frac{2}{3}}}$. Then this condition cannot be satisfied. Hence Born approximation cannot be used for scattering of slow electrons by atoms.

High energy: In the high energy limit, $\frac{k}{\varsigma} >> 1$, condition (14.81) becomes (in this case, the first term dominates)

$$\frac{\mu |V_0|}{\hbar^2 k \varsigma} \ln \left(\frac{2k}{\varsigma} \right) \ll 1 \qquad \text{or} \qquad \frac{|V_0| a}{\hbar v} \ln(2ka) \ll 1, \tag{14.83}$$

where $a = \frac{1}{\varsigma}$ and $v = \frac{\hbar k}{\mu}$ is the relative speed.

For Coulomb potential limit, $\frac{V_0}{\varsigma} = V_0 a = ZZ'e^2$ and $V_0 \to 0,\ a \to \infty\ (\varsigma \to 0)$. Then superficially it appears that Eq. (14.83) cannot be satisfied for $\frac{k}{\varsigma} = ka >> 1$. However a glance at Eq. (14.76) shows that if k is large, ς plays no role in the denominator, except when $\theta \simeq 0$. Hence $2ka$ may be kept finite, but large (except for scattering in forward direction, i.e., $\theta = 0$) such that $\ln(2ka) \sim$ unity and Eq. (14.83) becomes,

$$\frac{ZZ'e^2}{\hbar v} \ll 1 \quad \text{or} \quad k \gg \frac{\mu Z Z' e^2}{\hbar^2}. \tag{14.84}$$

This can be true at sufficiently high energy, which agrees with our general conclusions [see Ref. Merzbacher (1965)].

Direct application to Coulomb scattering

For Coulomb potential, $V(r) = \frac{ZZ'e^2}{r}$, and Eq. (14.78) becomes

$$\frac{2\mu}{\hbar^2 k} Z Z' e^2 \left| \int_0^\infty \frac{1}{r'} e^{ikr'} \sin(kr') dr' \right| \ll 1.$$

In high energy limit (k very large), the integrand oscillates rapidly and the integral decreases rapidly as k (energy) increases. Additional factor of $\frac{1}{k}$ makes left side smaller and the condition is satisfied. Hence Born approximation is valid in high energy limit.

The validity criterion at low energy ($k \to 0$), Eq. (14.80) becomes

$$ZZ'e^2 \int_0^\infty \frac{1}{r'} r' dr' \ll \frac{\hbar^2}{2\mu}.$$

The integral on left diverges. Hence Born approximation is not applicable in the low energy limit. However, note that our criterion (14.80) was obtained by replacing $\sin(kr')$ by kr' for $k \to 0$ in general, But this cannot be applied rigorously for the long-ranged Coulomb potential, since even for very small k, values of $r' \to \infty$ are needed (as $V(r')$ does not decrease faster than $\frac{1}{r'}$). We can get an estimate for the lower limit of energy in the following manner. We follow the argument leading to Eq. (14.80), but stop at Eq. (14.79) and then integrate directly, to get from Eq. (14.78)

$$\frac{2\mu}{\hbar^2 k} ZZ'e^2 \int_0^\infty \frac{\sin(kr')}{r'} dr' \ll 1$$

$$\text{i.e.,} \quad \frac{2\mu}{\hbar^2 k} ZZ'e^2 . \frac{\pi}{2} \ll 1,$$

since $\int_0^\infty \frac{\sin(kr')}{r'} dr' = \int_0^\infty \frac{\sin x}{x} dx = \frac{\pi}{2}$ $(k > 0)$. Hence

$$k \gg \frac{\mu ZZ'e^2 \pi}{\hbar^2} \tag{14.85}$$

is the criterion for validity of Born approximation for Coulomb potential. It gives us another estimate of the lower energy limit above which Born approximation is valid. This is about three times larger than the estimate (14.84)

[Comment: Note that the above treatment is not completely rigorous, since we used

$$|V(r')e^{ikr'} \sin(kr')| = |V(r')|\,|\sin(kr')|,$$

but in the integrand, replaced $|\sin(kr')|$ by $\sin(kr')$. This may be roughly justified as follows. For sufficiently small k, the wave length of $\sin(kr')$ is very large. Hence the amplitude of $\frac{\sin(kr')}{r'}$ decreases rapidly beyond the principal peak, for which the

sine function is positive. Thus the contribution to the integral comes mainly from the principal peak.]

To get a numerical estimate from Eq. (14.85), consider scattering of a proton (of mass M, nucleon mass, which is 938 MeV) by another proton (p-p scattering). Hence we have $\mu = \frac{M}{2}$, $Z = Z' = 1$. Therefore

$$\begin{aligned} E = \frac{\hbar^2 k^2}{2\mu} &\gg \frac{\mu}{2\hbar^2}(ZZ'e^2)^2\pi^2 \\ &= \frac{\frac{M}{2}}{2\hbar^2}(e^2)^2\pi^2 = \frac{Mc^2}{4}\left(\frac{e^2}{\hbar c}\right)^2\pi^2 \\ &\simeq \frac{938\ \text{MeV}}{4}.\left(\frac{1}{137}\right)^2\pi^2 \simeq\ 0.123\ \text{MeV}, \end{aligned}$$

where we used the fine structure constant $\alpha = \frac{e^2}{\hbar c} = \frac{1}{137}$. Thus at energies much above 123 KeV the Born approximation is reliable for p-p scattering.

14.9 Resonance Scattering

We saw in Chap. 8, Sect. 8.2 that if there is a finite barrier outside which a particle can be free, then this barrier can temporarily trap the particle in a quasi-bound state with energy (E_r) less than the barrier height (V_0). Suppose a three-dimensional spherically symmetric potential has such a barrier. Then an incident particle may be trapped in this barrier, and there will be a delay in the re-emergence of the particle, if its energy is equal to the energy of the quasi-bound state E_r for orbital angular momentum l. Then the total cross-section for the lth partial wave ($= \sigma_l$, integrated over all angles) will pass rapidly through a peak with a maximum of $\sigma_l = \frac{4\pi(2l+1)}{k^2}$, over its otherwise gradual variation with energy. Equation (14.27) shows that this maximum is attained when corresponding phase shift passes rapidly through $\delta_l = \frac{\pi}{2}$. Appearance of such a peak in the cross-section as a function of energy (or k) is called a resonance scattering. This phenomenon can be understood as follows. A particle with energy $E \neq E_r$ is scattered in the usual manner. However as $E \to E_r$, there is a large probability of the particle being trapped temporarily as a quasi-bound state. This increases the scattering amplitude and cross-section.

The centrifugal potential $V_{CF}(r) = \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}$ produces a convenient barrier for a neutron approaching a nucleus of radius R. As a simple model one can take the nuclear potential $V_N(r)$ to be a spherically symmetric square well of depth V_0.

$$\begin{aligned} V_N(r) &= -V_0 \quad \text{for } 0 \le r \le R \\ &= \ 0 \quad \text{for } r > R. \end{aligned}$$

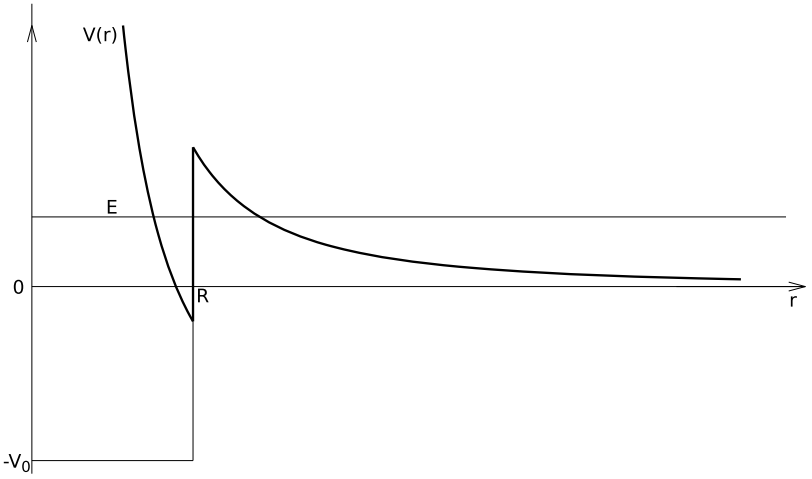


Fig. 14.14 A simple model potential for the resonance scattering of a neutron by a nucleus. The model potential consists of a square well of depth $-V_0$ and radius R plus the centrifugal repulsion for orbital angular momentum l. The complete potential is shown by the thick curve (for this plot l is taken to be 4). An incoming neutron with energy less than the height of the barrier, i.e., $E < \frac{\hbar^2}{2m}\frac{l(l+1)}{R^2}$, can have resonance scattering if its energy matches that of a quasi-bound state

Then the complete potential for the approaching neutron is $V(r) = V_N(r) + V_{CF}(r)$ (as shown in Fig. 14.14). It can trap the neutron in a quasi-bound state with $E > 0$. Corresponding σ_l then have resonance at an energy, close to the energy of a quasi-bound state. Plots of σ_l and δ_l showing resonance behavior can be found for a numerical calculation in Ref. Sakurai (2000).

For a spherically symmetric potential, partial wave expansion of the scattering amplitude is given by Eq. (14.25). Let us call the angle-independent part of it for the lth partial wave as $f_l(k)$, such that

$$f_l(k) = \frac{2l+1}{k} e^{i\delta_l} \sin\delta_l .$$

Now

$$\begin{aligned} e^{i\delta_l} \sin\delta_l &= (\cos\delta_l + i \sin\delta_l) \sin\delta_l = \frac{\cot\delta_l + i}{\mathrm{cosec}^2\delta_l} \\ &= \frac{\cot\delta_l + i}{1 + \cot^2\delta_l} = \frac{1}{\cot\delta_l - i} . \end{aligned} \tag{14.86}$$

As E crosses E_r [at which σ_l has a local maximum, see Eq. (14.27)] from below, δ_l passes through $\frac{\pi}{2}$ (or $\frac{3\pi}{2}, \cdots$) and $\cot\delta_l$ passes from positive to negative through zero [Sakurai (2000)]. For E close to E_r, we expand $\cot\delta_l$ as a function of E in a Taylor series about E_r

$$\cot\delta_l(E) = \cot\delta_l\Big|_{E=E_r} + \frac{d\cot\delta_l}{dE}\Big|_{E=E_r}(E-E_r)+\cdots$$
$$\approx 0 - c(E-E_r) \equiv -\frac{2}{\Gamma}(E-E_r), \tag{14.87}$$

where $c = -\frac{d\cot\delta_l}{dE}\Big|_{E=E_r} \equiv \frac{2}{\Gamma}$ is a positive quantity. Hence Γ is positive. Using the substitution of Eq. (14.87) in Eq. (14.86), the cross-section (integrated over all angles) [see Eqs. (14.25) and (14.27)] for the lth partial wave is

$$\sigma_l(E) = \frac{4\pi(2l+1)}{k^2}\left|e^{i\delta_l}\sin\delta_l\right|^2 = \frac{4\pi(2l+1)}{k^2}\left|\frac{1}{-\frac{2}{\Gamma}(E-E_r)-i}\right|^2$$
$$= \frac{4\pi(2l+1)}{k^2}\frac{\frac{\Gamma^2}{4}}{(E-E_r)^2+\frac{\Gamma^2}{4}}. \tag{14.88}$$

This equation shows that the partial cross-section as a function of incident energy E will have a peak at $E = E_r$ [see Fig. 7.12 of Ref. Sakurai (2000)]. At the peak, it reaches the maximum possible value $\sigma_l^{max} = \frac{4\pi(2l+1)}{k^2}$. Equation (14.88) is called the one level Breit–Wigner formula. Then $\sigma_l(E = E_r \pm \frac{\Gamma}{2}) = \frac{1}{2}\sigma_l^{max}$. It shows that Γ (its unit is that of energy) is the full width at half maximum of the peak. Γ is called the *width of the resonance*. The resonance peak is sharp for small Γ and broad for large Γ. When Γ is very large (extremely broad resonance, i.e., no resonance in the limit $\Gamma \to \infty$) the last factor in Eq. (14.88) approaches 1 and we have a monotonic energy dependent partial cross-section.

14.10 Problems

1. Derive Eq. (14.10) from Eq. (14.5).
2. Consider elastic scattering.
   (a) Investigate using Eq. (14.5) the range of θ_{lab} and θ_{CM} and their relation for $\gamma << 1$, $\gamma = 1$ and $\gamma >> 1$. Using a suitable numerical technique (e.g. gnu plot) plot θ_{lab} as a function of θ_{CM} for: $\gamma << 1$, $\gamma = 1$ and $\gamma >> 1$ cases.
   (b) Using Eq. (14.10) investigate the relation between $\left(\frac{d\sigma}{d\Omega}\right)_{\text{lab}}$ and $\left(\frac{d\sigma}{d\Omega}\right)_{\text{CM}}$ for $\gamma << 1$, $\gamma = 1$ and $\gamma >> 1$. Plot the factor in front of $\left(\frac{d\sigma}{d\Omega}\right)_{\text{CM}}$ in Eq. (14.10) as a function of θ_{CM} and comment on its dependence on θ_{CM}.
3. Consider three-dimensional elastic scattering of a quantum mechanical particle by a spherically symmetric potential

$$V(\vec{r}) = V_0 \qquad \text{for } r \le a$$
$$= 0 \qquad \text{for } r > a,$$

where V_0 is positive.

(a) Calculate the total cross-section for the $l = 0$ partial wave.
(b) Obtain an expression for the phase shift of the lth partial wave.
(c) What happens when $V_0 < 0$?

4. Calculate the Born amplitude and corresponding differential scattering cross-section for the potential of the previous problem. Plot these as functions of scattering angle.

References

Arfken, G.: Mathematical Methods for Physicists. Academic Press, New York (1966)
Merzbacher, E.: Quantum Mechanics. John Wiley & Sons Inc., New York (1965)
Sakurai, J.J. (ed. by Tuan, S.F.): Modern Quantum Mechanics. Addison-Wesley, Delhi (Second Indian reprint) (2000)
Schiff, L.I.: Quantum Mechanics, 3rd edn. McGraw-Hill Book Company Inc., Singapore (1968)
Sengupta, S.: Lecture Notes (2003) (unpublished)

Appendix
Orthogonality: Physical and Mathematical

Abstract In this appendix to Chap. 7, we distinguish between orthogonality relation arising from the postulates of quantum mechanics and that appearing in corresponding standard mathematical eigen value equation (where exact analytic solutions are possible). We show that in most cases they are consistent. Also the mathematical interval (required for hermiticity of the differential operator) agrees with the physical interval defined by the potential. These show close connections between physics and mathematics.

A number of standard quantum mechanical problems are solved in coordinate space by reducing them to standard mathematical equations. Usually the quantum mechanical (i.e. physical) orthogonality relation agrees with the orthogonality relation satisfied by the corresponding mathematical eigen value equation. Moreover the physical domain and mathematical interval [for requirement of hermiticity, see discussion after Eq. (7.26)] also agree. This shows a close connection between mathematics and physics. It may seem automatic. That this connection is not automatic is demonstrated by its failure for the hydrogen atom case, as we see in Ex. 5b below. In the following, we first briefly recapitulate steps usually taken to solve a typical case. Then we present a number of examples and a counter-example.

For a multi-dimensional problem, we can choose a radial (or hyper-radial) variable ρ and the remaining variables are collectively denoted by Ω. These are the physical variables. The Schrödinger equation for the eigen state λ reads

$$\hat{H}(\rho, \Omega)\psi_\lambda(\rho, \Omega) = E_\lambda \psi_\lambda(\rho, \Omega).$$

In a number of cases with certain symmetries, the variables ρ and Ω separate easily. In such cases, writing $\psi_\lambda(\rho, \Omega) = U_\lambda(\rho)\Phi_\lambda(\Omega)$ we get after separation of variables, a one-dimensional Schrödinger equation in ρ [with possible contributions in the effective Hamiltonian $\hat{h}(\rho)$ arising from other variables Ω]

$$\hat{h}(\rho)U_\lambda(\rho) = \lambda U_\lambda(\rho). \tag{A.1}$$

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5

This differential equation is usually solved by transforming the variable to $x = f(\rho)$, with a specified function $f(\rho)$ and writing

$$U_\lambda(\rho) = u_\lambda(x)g(x), \qquad \text{where } x = f(\rho). \tag{A.2}$$

The physical interval (determined by the potential) is $a \le \rho \le b$. The function $g(x)$ is also specified, usually arising from the asymptotic behavior of the one-dimensional Schrödinger equation (see Sect. 9.3, Chap. 9). Substitution of Eq. (A.2) in Eq. (A.1) gives a second order differential equation for $u_\lambda(x)$. This equation can be put in self-adjoint Sturm–Liouville (SL) form for its differential operator $\hat{\mathcal{L}}(x)$, which satisfies an eigen value equation

$$\hat{\mathcal{L}}(x)u_\lambda(x) = \lambda w(x)u_\lambda(x). \tag{A.3}$$

Sometimes the differential operator $\hat{\mathcal{L}}(x)$ can be identified as the differential operator for a standard mathematical function. In such cases, an exact analytic solution is possible. Quantum mechanical requirement that $\hat{\mathcal{L}}(x)$ must be Hermitian is usually satisfied if zeros (say a' and b') of its $p(x)$ [see Eq. (7.21)] are the mathematical boundaries. Hence the mathematical interval is $a' \le x \le b'$. [see Chap. 7, Sect. 7.1.4; specially Eq. (7.26) and the discussion following it]. In Eq. (A.3) $w(x)$ is the weight function appropriate for hermiticity requirement of $\hat{\mathcal{L}}(x)$, and λ is the eigen value. The set of eigen functions $\{u_\lambda(x)\}$ satisfies a *mathematical orthogonality relation*

$$\int_{a'}^{b'} u^*_{\lambda'}(x)u_\lambda(x)w(x)\mathrm{d}x = 0 \qquad \text{for } \lambda' \neq \lambda. \tag{A.4}$$

Requirement of *orthogonality of the physical solution* $U_\lambda(\rho)$ is

$$\int_{a}^{b} U^*_{\lambda'}(\rho)U_\lambda(\rho)\xi(\rho)\mathrm{d}\rho = 0 \qquad \text{for } \lambda' \neq \lambda, \tag{A.5}$$

where $\xi(\rho)$ is the phase space factor for ρ arising from the volume element $\mathrm{d}\tau = \xi(\rho)\mathrm{d}\rho\mathrm{d}\Omega$. For example, for spherical polar coordinates (r, θ, ϕ), we have $\mathrm{d}\tau = r^2\mathrm{d}r\sin\theta\mathrm{d}\theta\mathrm{d}\phi = r^2\mathrm{d}r\mathrm{d}\Omega$. Hence taking $\rho = r$, we have $\xi(\rho) = \rho^2$. Substituting Eq. (A.2) in Eq. (A.5) and transforming the variable of integration from ρ to x, we have a *physical orthogonality relation*

$$\int_{a_x}^{b_x} u^*_{\lambda'}(x)u_\lambda(x)|g(x)|^2\tilde{\xi}(x)\mathrm{d}x = 0 \qquad \text{for } \lambda' \neq \lambda, \tag{A.6}$$

where $a_x = f(a)$, $b_x = f(b)$ and $\tilde{\xi}(x)$ comes from the transformation of the variable from ρ to x, including the function $\xi(\rho)$. In general it is found that $a' = a_x$, $b' = b_x$, i.e., the mathematical and physical intervals agree. We also usually have $w(x) = |g(x)|^2\tilde{\xi}(x)$. Thus the mathematical orthogonality relation (A.4) agrees with the physical orthogonality relation (A.6). This demonstrates a close connection between mathematics and physics. We present a few examples and a counter-example below.

Ex. 1. Eigen Value Equations for Legendre Equations

Standard form of associated Legendre equation (7.45) is

$$(1-x^2)y''(x) - 2xy'(x) + \left[l(l+1) - \frac{m^2}{1-x^2}\right]y(x) = 0 \qquad (l, m \text{ are real constants}). \tag{A.7}$$

Legendre equation is a special case for $m = 0$. Both equations are already in SL form with $p(x) = (1-x^2)$. This vanishes for $x = \pm 1$. Hence $a' = -1, b' = 1$ and mathematical interval is $-1 \le x \le 1$. Both equations have eigen value $\lambda = l(l+1)$ and $w(x) = 1$. For $l =$ a non-negative integer, the standard regular (i.e. analytic in the entire interval) solution is $y(x) = P_l^m(x)$. Then the mathematical orthogonality relation is

$$\int_{-1}^{1} P_{l'}^m(x) P_l^m(x) \mathrm{d}x = 0 \quad \text{for } l' \neq l. \tag{A.8}$$

This relation is also valid for the Legendre function $P_l(x)$, for which $m = 0$.

Comparison with a Physical Case: Eigen Value Equation of Orbital Angular Momentum Squared ($\hat{L}^2$) Operator

Eigen function of the physical $\hat{L}^2$ operator (Chap. 10) is $\Theta_l^m(\theta)$ satisfying the orthogonality relation in physical (θ) space (for m fixed)

$$\int_0^{\pi} \Theta_{l'}^m(\theta)\, \Theta_l^m(\theta) \sin\theta \mathrm{d}\theta = 0 \qquad \text{for } l' \neq l. \tag{A.9}$$

Here the phase space factor $\xi(\theta) = \sin\theta$. Changing the polar angle variable (θ) to $x = \cos\theta$ we have $f(\theta) = \cos\theta$ and $\sin\theta \mathrm{d}\theta = -\mathrm{d}x$. Also $\Theta_l^m(\theta) = P_l^m(x)$.

Hence $g(x) = 1$ and $\tilde{\xi}(x) = 1$. The interval for θ is $[0, \pi]$, and that for x is $[\cos(\pi), \cos(0)] = [-1, 1]$. Then we see that Eq. (A.9) agrees with Eq. (A.8). Thus the mathematical interval and orthogonality relation agree with corresponding physical quantities.

Ex. 2. Eigen Value Equation for Hermite Differential Equation

Standard Hermite equation (7.37) is

$$y''(x) - 2xy'(x) + 2ay(x) = 0 \quad (a \text{ is a real constant}). \tag{A.10}$$

It is not in SL form. Multiplying with e^{-x^2} puts it in that form

$$\mathrm{e}^{-x^2} y''(x) - 2x\,\mathrm{e}^{-x^2} y'(x) + 2a\,\mathrm{e}^{-x^2} y(x) = 0.$$

We identify $p(x) = \mathrm{e}^{-x^2}$, $w(x) = \mathrm{e}^{-x^2}$ and $\lambda = 2a$. Hence $p(x) = 0$ for $x \to \mp\infty$ and mathematical interval is $-\infty < x < \infty$. For $a = n$ (a non-negative integer) the standard regular solution is the Hermite polynomial $H_n(x)$, which satisfies a mathematical orhogonality relation, involving the weight function e^{-x^2} as follows

$$\int_{-\infty}^{\infty} H_{n'}(x) H_n(x) \mathrm{e}^{-x^2} \mathrm{d}x = 0 \qquad \text{for } n' \neq n. \tag{A.11}$$

Comparison with a Physical Case: 1D Harmonic Oscillator (Sec. 9.4, Chap. 9)

The eigen function in the physical interval $-\infty < x < \infty$ is given by Eq. (9.27)

$$u_n(x) = N_n H_n(\alpha x)\, \mathrm{e}^{-\frac{1}{2}\alpha^2 x^2}, \tag{A.12}$$

where α is a constant. The physical orthogonality relation is

$$\int_{-\infty}^{\infty} u_{n'}(x)\, u_n(x) \mathrm{d}x = 0 \qquad \text{for } n' \neq n.$$

Substituting $u_n(x)$ from Eq. (A.12) and changing the variable to $\zeta = \alpha x$, we have (note that $g(\zeta) = \mathrm{e}^{-\frac{1}{2}\zeta^2}$ and $\tilde{\xi}(\zeta) = 1$)

$$\int_{-\infty}^{\infty} H_{n'}(\zeta) H_n(\zeta) \mathrm{e}^{-\zeta^2} \mathrm{d}\zeta = 0 \qquad \text{for } n' \neq n. \tag{A.13}$$

In this case the mathematical and physical intervals are the same. Equations (A.11) and (A.13) show that their orthogonality relations are also the same.

Ex. 3. Eigen Value Equation for Bessel Differential Equation

Standard Bessel differential equation (7.56) is

$$x^2 y''(x) + x y'(x) + (x^2 - \nu^2) y(x) = 0 \quad (\nu \text{ is a real constant}). \tag{A.14}$$

Its standard regular solution is $y(x) = J_\nu(x)$, the Bessel function. For a fixed value of ν, this equation apparently does not have an eigen value term. However replacing x by $\frac{\zeta_\nu}{a} x$ (where ζ_ν and a are constants) and dividing by x we have

$$x \frac{d^2 J_\nu(\frac{\zeta_\nu}{a} x)}{\mathrm{d}x^2} + \frac{\mathrm{d} J_\nu(\frac{\zeta_\nu}{a} x)}{\mathrm{d}x} + x\Big(\frac{\zeta_\nu^2}{a^2} - \frac{\nu^2}{x^2}\Big) J_\nu\Big(\frac{\zeta_\nu}{a} x\Big) = 0. \tag{A.15}$$

Then we can identify the quantity $\frac{\zeta_\nu^2}{a^2}$ as the eigen value. This equation is in SL form with

$$p(x) = x, \quad q(x) = -\frac{\nu^2}{x}, \quad \lambda = \frac{\zeta_\nu^2}{a^2}, \quad \text{and} \quad w(x) = x.$$

We next have to make sure that the corresponding differential operator is Hermitian. Since $p(x) = 0$ has only one solution, *viz*,, $x = 0$, we can choose one extremity of the interval at $x = 0$. Then choosing the other extremity as $x = a$, the condition for hermiticity demands (Dirichlet boundary condition)

$$J_\nu\Big(\frac{\zeta_\nu}{a} x\Big)\Big|_{x=a} = 0, \quad \text{i.e.,} \quad J_\nu(\zeta_\nu) = 0. \tag{A.16}$$

Thus ζ_ν is a zero of the Bessel function $J_\nu(x)$. Let it be the nth zero $\zeta_{\nu,n}$. Then the eigen value of Eq. (A.15) is $\frac{\zeta_{\nu,n}^2}{a^2}$, with eigen value index n. The complete mathematical orthonormality relation is [see Eq. (5.46) of Ref. Chattopadhyay (2006)]

$$\int_0^a J_\nu\Big(\frac{\zeta_{\nu,n}}{a}x\Big)J_\nu\Big(\frac{\zeta_{\nu,m}}{a}x\Big)\,x\,\mathrm{d}x = \frac{a^2}{2}\big[J_{\nu+1}(\zeta_{\nu,n})\big]^2\,\delta_{n,m}\,, \tag{A.17}$$

where the factor x in the integrand comes from the weight function, see Eq. (A.4).

Comparison with a Physical Case: A Particle in a Cylindrical Box (Sect. 12.3, Chap. 12)

Consider a quantum mechanical particle in a cylindrical box (of radius a and height H) with rigid walls. The cylindrical variables are (r, ϕ, z) with the physical intervals

$$0 \le r \le a, \qquad 0 \le \phi \le 2\pi, \qquad 0 \le z \le H.$$

Volume element is

$$\mathrm{d}V = r\mathrm{d}r\mathrm{d}\phi\mathrm{d}z,$$

so that the phase space factor in r-space is $\xi(r) = r$. Full wave function is given by Eq. (12.16)

$$u_{p,m,n}(r, \phi, z) = DJ_m\Big(\frac{\zeta_{m,p}}{a}r\Big)\sin\Big(\frac{n\pi z}{H}\Big)\mathrm{e}^{im\phi}, \tag{A.18}$$

whose radial part is just $J_m\big(\frac{\zeta_{m,p}}{a}r\big)$, without any r-dependent additional factor, i.e., $g(r) = 1$. Related quantum number is p. Orthogonality relation for the radial part (including the phase space factor $\xi(r)$) is

$$\int_0^a J_m\Big(\frac{\zeta_{m,p}}{a}r\Big)J_m\Big(\frac{\zeta_{m,p'}}{a}r\Big)r\mathrm{d}r = 0 \qquad \text{for } p' \ne p. \tag{A.19}$$

The physical orthogonality relation (A.19) is just the same as the mathematical orthogonality relation (A.17), with appropriate changes.

Ex. 4. Eigen Value Equation for Spherical Bessel Functions

Standard form of spherical Bessel equation is [Eq. (7.69)]

$$x^2y''(x) + 2xy'(x) + [x^2 - l(l+1)]y(x) = 0,$$
$$(l \text{ is an integral constant } \ge 0). \tag{A.20}$$

This equation is already in SL form, but there is no clear eigen value term for a fixed l. Substituting

$$x = \frac{z_{l,m}}{a}\rho, \quad \text{where} \quad z_{l,m} \quad \text{and} \quad a \quad \text{are constants,}$$

in Eq. (A.20) and writing $y(x) = j_l\left(\frac{z_{l,m}}{a}\rho\right)$ (where $j_l(x)$ ia the standard regular solution, called spherical Bessel function), we get

$$\rho^2 \frac{d^2}{\mathrm{d}\rho^2} j_l\left(\frac{z_{l,m}}{a}\rho\right) + 2\rho \frac{d}{\mathrm{d}\rho} j_l\left(\frac{z_{l,m}}{a}\rho\right) + \left[\frac{(z_{l,m})^2}{a^2}\rho^2 - l(l+1)\right] j_l\left(\frac{z_{l,m}}{a}\rho\right) = 0. \quad \text{(A.21)}$$

Now Eq. (A.21) is an eigen value equation in SL form, for which we identify

$$p(\rho) = \rho^2, \quad q(\rho) = -l(l+1), \quad \lambda = \frac{(z_{l,m})^2}{a^2}, \quad \text{and} \quad w(\rho) = \rho^2.$$

Hence, $p(\rho) = 0$ gives $\rho = 0$ as one of the extremities of the interval. For the other we choose $\rho = a$ (so that mathematical interval is $0 \le \rho \le a$), together with the additional boundary condition for hermiticity (Dirichlet boundary condition)

$$j_l\left(\frac{z_{l,m}}{a}\rho\right)\Big|_{\rho=a} = j_l\left(z_{l,m}\right) = 0. \quad \text{(A.22)}$$

Thus $z_{l,m}$ is the mth zero of $j_l(x)$.
The orthonormality relation is [see Eq. (5.99) of Ref. Chattopadhyay (2006)]

$$\int_0^a j_l\left(\frac{z_{l,m}}{a}\rho\right) j_l\left(\frac{z_{l,n}}{a}\rho\right)\rho^2 \mathrm{d}\rho = \frac{a^3}{2}\left[j_{l+1}\left(z_{l,m}\right)\right]^2 \delta_{m,n}. \quad \text{(A.23)}$$

Note that the factor ρ^2 in the integrand comes from the weight function and the mathematical interval is $0 \le \rho \le a$.

Comparison with a Physical Case: A Particle in a Spherical Hole

Consider a particle in a spherical hole with rigid walls having radius a and centered at the origin (Sect. 12.1, Chap. 12). In spherical polar coordinates (r, θ, ϕ), the volume element is $r^2 \mathrm{d}r \sin\theta \mathrm{d}\theta \mathrm{d}\phi = r^2 \mathrm{d}r \mathrm{d}\Omega$. Hence the phase space factor $\xi(r) = r^2$. The radial wave function is given by Eq. (12.6)

$$R_{n,l}(r) = A_{n,l}\, j_l\Big(\frac{z_{l,n}}{a} r\Big),$$

where $z_{l,n}$ is the nth zero of $j_l(x)$. The physical orthogonality relation of the radial wave functions is given by (note that phase space factor is r^2 and physical interval is $0 \le r \le a$)

$$\int_0^a j_l\Big(\frac{z_{l,n}}{a} r\Big) j_l\Big(\frac{z_{l,n'}}{a} r\Big) r^2 \mathrm{d}r = 0 \qquad \text{for } n' \neq n. \tag{A.24}$$

This agrees with the mathematical orthogonality relation (A.23), with obvious change of variable and quantum numbers.

Ex. 5. Eigen Value Equation for Associated Laguerre Functions

Standard form of associated Laguerre equation is given by Eq. (7.52), in which we replaced constants p, q by P, Q to avoid confusion with $p(x), q(x)$ respectively

$$xy''(x) + (P + 1 - x)y'(x) + (Q - P)y(x) = 0, \quad (P, Q \text{ constants}). \tag{A.25}$$

It can be put in SL form by multiplying through by $x^P \mathrm{e}^{-x}$ to get the eigen value equation for $y(x) \equiv L_Q^P(x)$ (called associated Laguerre function) as

$$\frac{d}{\mathrm{d}x}\Big[x^{P+1}\mathrm{e}^{-x}\frac{d}{\mathrm{d}x}\Big]L_Q^P(x) + (Q - P)x^P \mathrm{e}^{-x} L_Q^P(x) = 0.$$

Comparing this with the standard eigen value equation (7.22), together with the SL form of the differential operator, Eq. (7.21), we identify

$$p(x) = x^{P+1}\mathrm{e}^{-x}, \quad q(x) = 0, \quad \lambda = (Q - P), \quad w(x) = x^P \mathrm{e}^{-x}, \quad u_\lambda(x) = L_Q^P(x).$$

The eigen value actually depends on Q, since P is a constant for the equation, appearing in $p(x)$. As $p(x)$ vanishes at $x = 0$ and $x \to \infty$, the mathematical interval is $0 \le x < \infty$. The eigen function $L_Q^P(x)$ is a regular solution and becomes a polynomial of degree $(Q - P)$, if $(Q - P)$ is a non-negative integer. The mathematical orthogonality relation (with weight function $x^P \mathrm{e}^{-x}$ in the interval $0 \le x < \infty$) is [see Ref. Chattopadhyay (2006)]

$$\int_0^\infty L_{Q'}^P(x) L_Q^P(x) x^P \mathrm{e}^{-x} \mathrm{d}x = 0 \qquad \text{for } (Q' - P) \neq (Q - P). \tag{A.26}$$

Comparison with Physical Cases

(a) 3-D Harmonic Oscillator (Sect. 12.4, Chap. 12)

The radial wave function is given by Eq. (12.26)

$$R_{n,l}(r) = \frac{\tilde{R}_{n,l}(r)}{r} \propto (\alpha r)^l \, L^{l+\frac{1}{2}}_{\frac{\lambda}{4}+\frac{l}{2}-\frac{1}{4}}(\alpha^2 r^2) \, \mathrm{e}^{-\frac{1}{2}\alpha^2 r^2} \propto (\alpha r)^l \, L^{l+\frac{1}{2}}_{n+l-\frac{1}{2}}(\alpha^2 r^2) \, \mathrm{e}^{-\frac{1}{2}\alpha^2 r^2},$$

where $\alpha = \sqrt{\frac{\mu\omega_c}{\hbar}}$ is a constant (note that it is also independent of the quantum number n). The last proportionality is obtained by Eq. (12.25), the effective radial quantum number being n. The physical orthogonality relation (note that phase space factor is r^2 and physical interval is $0 \le r < \infty$) is

$$\int_0^\infty R_{n',l}(r) \, R_{n,l}(r) \, r^2 \mathrm{d}r = 0 \qquad \text{for } n' \neq n. \tag{A.27}$$

Using $R_{n,l}(r)$ from above, left hand side of Eq. (A.27) is

$$\text{LHS} \propto \int_0^\infty \left[L^{l+\frac{1}{2}}_{n'+l-\frac{1}{2}}(\alpha^2 r^2) \, L^{l+\frac{1}{2}}_{n+l-\frac{1}{2}}(\alpha^2 r^2) \right] (\alpha r)^{2l} \, \mathrm{e}^{-\alpha^2 r^2} \, r^2 \mathrm{d}r.$$

Putting $\zeta = \alpha^2 r^2$,

$$\text{LHS} \propto \int_0^\infty L^{l+\frac{1}{2}}_{n'+l-\frac{1}{2}}(\zeta) \, L^{l+\frac{1}{2}}_{n+l-\frac{1}{2}}(\zeta) \, \zeta^{l+\frac{1}{2}} \, \mathrm{e}^{-\zeta} \, \mathrm{d}\zeta = 0 \quad \text{for } n' \neq n, \tag{A.28}$$

the last condition is by Eq. (A.27). This physical orthogonality relation agrees exactly with the mathematical orthogonality relation, Eq. (A.26), with $P = l + \frac{1}{2}$, $Q = n + l - \frac{1}{2}$ and $Q' = n' + l - \frac{1}{2}$.

(b) Hydrogen atom (Sect. 11.2, Chap. 11)

The radial wave function [note that $u_{n,l,m}(\vec{r}) = R_{n,l}(r) Y_{l,m}(\theta, \phi)$] is given by Eq. (11.17)

$$R_{n,l}(r) \propto \mathrm{e}^{-\frac{1}{2}\rho} \rho^l L^{2l+1}_{n+l}(\rho), \tag{A.29}$$

where $\rho = \alpha r$ and $\alpha = \frac{2Z}{n a_{\text{Bohr}}}$. The physical orthogonality relation is (note again that phase space factor is r^2, and physical interval is $0 \le r < \infty$ for the physical variable r)

$$\int_0^\infty R_{n',l}(r)\,R_{n,l}(r)\,r^2\mathrm{d}r = 0 \qquad \text{for } n' \neq n. \tag{A.30}$$

Using Eq. (A.29) and putting $\rho = \alpha r$, left side of Eq. (A.30) is

$$\text{LHS} \propto \int_0^\infty L_{n'+l}^{2l+1}(\rho) L_{n+l}^{2l+1}(\rho)\ \rho^{2l+2}\,\mathrm{e}^{-\rho}\,\mathrm{d}\rho. \tag{A.31}$$

This *does not agree with corresponding term of the mathematical orthogonality relation*, Eq. (A.26), with $P = 2l + 1$, $Q = n + l$ and $Q' = n' + l$. However, we note that although α is a constant (so that $\rho \propto r$), the constant of proportionality α involves the quantum number n $\left(viz,\ \alpha = \frac{2Z}{n\,a_{\text{Bohr}}}\right)$ [see Eq. (11.17)]. In contrast, for the 3-D harmonic oscillator the constant (α) involved in the transformation of the physical variable r to ζ is a pure constant independent of the quantum number.

Reference

Chattopadhyay, P.K.: Mathematical Physics. New Age International (P) Ltd., New Delhi (2006)

Index

A
Associated Laguerre function, 219, 236
Associated Legendre equation, 209

B
Basis, 18
  change of, 33
  closure relation, 32, 39
Bessel function, 140
Black body radiation, 4
Bohr-Sommerfeld quantization, 164
Born approximation, 289
  validity, 291
Bound state
  estimate number of, 164
  general procedure, 188
  quantum numbers, 163
    choice of, 165
    good, 165
Boundary conditions, 149
  and eigen values, 153
    continuous, 160
    discrete, 156
  at infinite potential step, 151
Boundary value problem, 122
Breit-Wigner formula, 298

C
Cauchy sequence, 40
Classical physics, 1
Classical turning points, 154
Complementarity principle, 7
Complete orthonormal set, 25, 29, 166
Compton scattering, 5
Continuity condition, 150
Correspondence principle, 9, 177
Coulomb scattering, 277
Cylindrical hole
  rigid walls, 230

D
De Broglie relation, 6
Degeneracy, 167
  3D harmonic oscillator, 233, 236
  relation to symmetry, 167
Delta function, 62
Differential equation
  associated Laguerre, 219
  associated Legendre, 209
  Bessel, 140
  Hermite, 132, 191
  Laguerre, 133
  Legendre, 131
  series solution, 117
  simple harmonic, 134
  singularity, 114
  spherical Bessel, 143, 226
  spherical modified Bessel, 145, 229
Diffraction experiment, 7
Dirac
  bra ket notation, 36
  delta function, 62
Discrete values of physical quantities, 6, 8
Dual nature, 2, 5, 6

E
Ehrenfest theorem, 175
Eigen functions, 9
Eigen values, 9
Einstein, 5

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

T. K. Das, *Quantum Mechanics*, UNITEXT for Physics,
https://doi.org/10.1007/978-981-99-0494-5

Endothermic process, 255
Equation of continuity, 61
Exothermic process, 255
Expansion postulate, 52
Expectation value, 53

F
Fourier transform
  integral representation, 78
  inverse transform, 78
Frobenius method, 118
Fundamental postulates, 50

G
Generating function
  Hermite polynomial, 193
  Legendre polynomial, 210
Good quantum numbers, 165
Gram determinant, 21
Gram-Schmidt orthonormalization, 24
Green's function
  in scattering, 281
  solving inhomogeneous differential eqn., 283

H
Hankel function, 141
Harmonic oscillator, 189
  importance in physics, 99
  matrix elements, 105, 195
  by operator method, 100
  uncertainty relations, 108
  wave function, 106, 193
Heisenberg, 7
Hilbert space, 41
Hydrogen atom, 215

I
Importance of symmetry, 165
  and choice of variables, 167
Inner product space, 21

K
Kronecker delta, 23

L
Legendre polynomial, 210
  Rodrigues' formula, 210
Light, corpuscular nature, 4
Linear combination, 17
Linear dependence, 17
Linear independence
  of functions, 115
    test for, 116
  of vectors, 17
    test for, 21
Linear vector space, 13
  angle between vectors, 23
  basis, 18
  completeness, 40
  dimension, 18
  infinite dimensional, 39
  inner product, 20
  length of vector, 23
  operators, 25
  in quantum mechanics, 49
Lippmann-Schwinger equation, 50, 71, 285

M
Matrix mechanics, 10, 78
Modified Bessel function, 142
Momentum eigen function, 76

N
Neumann function, 140
Normalization, 65
  box, 66
  delta function, 67
  flux, 65
Norm of a vector, 23
N-tuple, 15
Number of bound states in a well, 164

O
Observables, 51
  compatible, 92
  complete set of commuting, 166
  eigen values, 51
  variance of, 92
Old quantum theory, 6
Operator, 25
  anti-linear, 27
  commutator, 28
  creation, annihilation, 100
  eigen value equation, 29
  Hermitian, 29
  Hermitian adjoint, 28
  Hermitian differential, 129
  integral transform, 26

inverse, 28
linear, 25
as matrices, 30
non-linear, 27
prescription for QM, 55
projection, 39
rotation, 26
Optical theorem, 264
Orbital angular momentum, 205

P
Parabolic coordinates, 277
Parity, 168
Partial wave expansion of plane wave, 262
Periodic boundary conditions, 66
Phase shift, 265
relation with potential nature, 267
Photo-electric effect, 5
Photons, 2, 4
Physical system, 51
Pictures
Heisenberg, 85
Interaction, 87
Schrödinger, 84
Planck-Einstein relations, 5
Planck's constant, 4
Potential
cylindrical well
rigid walls, 230
harmonic oscillator 1D, 189
harmonic oscillator 3D, 233
square well
finite 1D, 183
finite 3D, 228
infinite 1D, 179
infinite 3D, 225
with delta function, 198
Probabilistic interpretation, 52
Probability current density, 61
Probability density, 60

Q
Quantization
condition, 6
of energy, 55
first, 8
Quantization postulate, 53
Quantum dynamics, 83
Quasi-bound state, 162, 200

R
Ramsauer-Townsend effect, 272
Recurrence relations
Hermite polynomial, 194
Representations
change, 77
coordinate, 71
matrix, 78
momentum, 74
occupation number, 105
R.m.s deviation, 92

S
Scattering
one dimension, 239
delta barrier, 245
finite square barrier, 241
rigid wall, 239
three dimension
asymptotic wave function, 258
Born approximation, 289
Coulomb, 277
differential cross-section, 254, 255, 260
experimental setup, 250
Green's function, 281, 287
kinematics, 252
Lippmann-Schwinger equation, 285
optical theorem, 264
partial waves, 260
phase shift, 265
resonance scattering, 296
rigid sphere, 274
scattering amplitude, 258, 288
total cross-section, 260
three dimension, 249
Schrödinger equation, 9, 56
coordinate representation, 72
momentum space, 75
time dependent, 58, 61
time independent, 112
Schwarz inequality, 22
Second order differential equation, 113
singularity, 114
Spherical Bessel functions, 144, 226
Spherical Hankel function, 144
Spherical harmonics, 212
Spherical hole
permeable walls, 228
rigid walls, 225
Spherical modified Bessel functions, 229
Spherical Neumann function, 144

State of a system, 51
bound and unbound, 60
localized, 60, 149
quasi-bound, 162
stationary, 113
Sturm-Liouville theory, 122
Subspace, 16

T
Time evolution, 56
Tunneling, quantum mechanical, 162
Two-body with mutual force, 215

U
Uncertainty principle, 7
Uncertainty relation, 54, 94
derivation of, 91
minimum uncertainty state, 95

V
Virial theorem, 178

W
Wave function
coordinate space, 58
free particle, 169
nodes, 160, 161, 181
probabilistic interpretation, 59
Wave mechanics, 10
Wave packet motion
general, 170
in harmonic oscillator well, 197
Wronskian, 116

Z
Zero point energy, 161